

SANDIA REPORT

SAND2011-3079

Unlimited Release

Printed May 2011

Bayesian estimation of unresolved fine-scale properties of binary media from multiscale observations

J. Ray, S. A. McKenna, B. van Bloemen Waanders, and Y. M. Marzouk
Sandia National Laboratories, CA
Sandia National Laboratories, NM and
Massachusetts Institute of Technology, MA

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory managed and operated by Sandia Corporation,
a wholly owned subsidiary of Lockheed Martin Corporation, for the United States
Department of Energy's National Nuclear Security Administration under
Contract DE-AC04-94-AL85000.

Approved for public release; further dissemination unlimited.



Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: <http://www.doe.gov/bridge>

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online ordering: <http://www.ntis.gov/ordering.htm>



Bayesian estimation of unresolved fine-scale properties of binary media from multiscale observations

Jaideep Ray

Sandia National Laboratories, Livermore CA 94551

Sean A. McKenna, and Bart van Bloemen Waanders
Sandia National Laboratories, Albuquerque NM 87185

Youssef M. Marzouk

Massachusetts Institute of Technology, Cambridge, MA 02139
{jairay,samcken,bartv}@sandia.gov, ymarz@mit.edu

Abstract

We present a Bayesian technique to estimate the fine-scale properties of a binary medium from multiscale observations. The binary medium of interest consists of a low-permeability matrix with embedded high-permeability inclusions. The inclusions are far smaller than the domain sizes of interest, and thus are never explicitly resolved. We consider the problem of estimating the spatial distribution of the inclusion proportion, $\mathbf{F}(\mathbf{x})$, and a characteristic length-scale of the inclusions, δ , from sparse multiscale measurements. The observations consist of coarse-scale (of the order of the domain size) measurements of the effective permeability of the medium (i.e., static data) and tracer breakthrough times (i.e., dynamic data) at a sparsely distributed set of locations. This ill-posed problem is regularized by specifying a Gaussian process model for the unknown field $\mathbf{F}(\mathbf{x})$ and expressing it as a superposition of Karhunen-Loève modes. The effect of the fine-scale structures on the coarse-scale effective permeability i.e., upscaling, is performed using a subgrid-model which includes δ as one of its parameters. A statistical inverse problem is posed to infer the weights of the Karhunen-Loève modes and δ , which is then solved using an adaptive Markov Chain Monte Carlo method. The solution yields distributions for the objects of interest, thus providing most probable estimates and uncertainty bounds on latent fine-scale structures. The technique is tested using synthetic data. The individual contribution of the static and dynamic data to the inference are also analyzed.

Acknowledgment

This work was supported Sandia National Laboratories' LDRD (Laboratory Directed Research and Development) funds, sponsored by the Computational and Information Sciences IAT. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U. S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Contents

1	Introduction	13
2	Literature review	15
2.1	Estimation of random fields	15
2.2	Topological and upscaling models for random binary media	18
2.3	Adaptive Markov chain Monte Carlo techniques	20
3	Models used in the inverse problem	23
3.1	Karhunen-Loève expansions of random fields	23
3.2	The transport model $\mathcal{M}(\mathbf{K})$	24
3.3	The link function \mathcal{L}	25
4	Multiscale inference with homoscedastic errors	31
4.1	Posing the inverse problem	31
4.2	Solving the inverse problem	33
4.3	Results and discussion	35
4.4	Summary	39
5	Posterior predictive modeling	49
6	Multiscale inference with model errors	57
6.1	Construction of the error model for the link function	57
6.2	Adapting and solving the inverse problem for model errors	60
7	Conclusions	67
	References	75

Figures

1	(a) The 2D rectangular domain \mathcal{D} with associated dimensions. No-flow boundary conditions are defined on $\partial\mathcal{D}$. A grid-block showing the collocation for pressure p and the velocities (u, v) in the x- and y-directions is shown. (b) We show a schematic of a binary medium with the light, higher permeability inclusions in a darker lower-permeability matrix. (c) The true $\mathbf{F}_c^t(\mathbf{x})$ field for the proportion of inclusions. (d) The true upscaled log-permeability field \mathbf{K}_c^t . (e) The 30×20 coarse-scale computational mesh with the locations of 20 sensors (sensor-set A, SSA). (f) We show the locations of the 34 sensors in sensor-set B (SSB).	29
2	Plots of \mathbf{K}_e calculated using the link function \mathcal{L} , as well as numerically estimated from a random binary medium using MODFLOW-2005, as a function of $\mathbf{F}(\mathbf{x})$. Left: The plots are for a medium where $\delta = 10$ grid-cells. Right: $\delta = 75$ grid-cells. We see that the numerically estimated \mathbf{K}_e form a cloud around the value predicted by \mathcal{L} . This is because $\{\mathbf{F}(\mathbf{x}), \delta\}$ is an incomplete description of a random binary medium, and an infinite number of realizations, each with its own permeability, can be conditioned on such an incomplete specification. The break in \mathcal{L} predictions at $\mathbf{F} \approx 0.5$ is due to percolation effects.	30

- 3 Plot of the joint and marginal probability densities of w_1, w_{15}, w_{30} and $\ln(\delta)$ from the posterior distribution. The joint distributions show very little correlations between the Karhunen-Loève modes at the large, medium and small scales, as well as their correlations with δ . Also, PDF of the individual parameters are roughly Gaussian, with the exception of δ 41
- 4 Results from the inference, computed using 9500 samples from the posterior distribution. In the top row, we plot the means $\bar{\mathbf{F}} = \mathbb{E}[\mathbf{F}(\mathbf{x})]$ and $\bar{\mathbf{K}}_e = \mathbb{E}[\mathbf{K}_e]$; they bear a strong resemblance to the true values in Fig. 1 (middle row). In the middle row, we plot the grid-block-wise standard deviation of $\mathbf{F}_i(\mathbf{x})$ and $\mathbf{K}_{e,i}$; they are smallest at the sensor locations. In the bottom row, we plot the errors $\epsilon_F = \mathbf{F}'_c(\mathbf{x}) - \bar{\mathbf{F}}$ and $\epsilon_K = \mathbf{K}'_c - \bar{\mathbf{K}}_e$, which show large values in regions of high gradient where the representation errors due to the smooth Gaussian process model are the largest. 42
- 5 Results from the inference, computed using 9500 samples from the posterior distribution. Only static data was used to draw the inferences. In the top row, we plot the means $\bar{\mathbf{F}}$ and $\bar{\mathbf{K}}_e$ which are worse than those seen in Fig. 4. Further, $\bar{\mathbf{K}}_e$ has larger errors than $\bar{\mathbf{F}}$ due to its finer structures. In the middle row, we plot the grid-block-wise standard deviation of $\mathbf{F}_i(\mathbf{x})$ and $\mathbf{K}_{e,i}$; they are larger and more widely distributed than their counterparts in Fig. 4. In the bottom row, we plot the errors ϵ_F and ϵ_K , which show larger values in regions of high gradient. Further, the errors are higher and more widely distributed compared to their counterparts in Fig. 4. 43
- 6 Results from the inference, computed using 9500 samples from the posterior distribution. Only dynamic data was used to draw the inferences. In the top row, we plot the means $\bar{\mathbf{F}}$ and $\bar{\mathbf{K}}_e$. The inferences are far worse than those seen in Fig. 4 and show little resemblance to the true $\mathbf{F}'_c(\mathbf{x})$ and \mathbf{K}'_c plotted as contours over them. Given such a large lack of fidelity, the standard deviations and errors plotted in the middle and last rows have little significance, but are reproduced here for completeness. Note that while Fig. 4 and Fig. 5 share the same color scales, the range had to be adjusted in this figure. 44

- 7 Marginalized PDFs of w_1, w_{15}, w_{30} and $\ln(\delta)$ as inferred from $\mathbf{k}^{(obs)}$ -only (dotted lines), $\mathbf{t}_b^{(obs)}$ -only (dashed lines) and jointly from $(\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)})$ (solid lines). The priors are plotted with ∇ for comparison. Top left: we see that the posterior for w_1 is almost entirely accounted for by the $\mathbf{k}^{(obs)}$ -only inversion (“static data only”); the posterior for w_1 , when inferred from $\mathbf{t}_b^{(obs)}$ only (“dynamic data only”) is little different from the prior. Top right: We see that both static and dynamic data contribute to the posterior of w_{15} ; the joint $(\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)})$ inversion is quite different from the other two. Bottom left: Surprisingly, neither $\mathbf{k}^{(obs)}$ nor $\mathbf{t}_b^{(obs)}$ are individually informative of the finer scales. However, when $\mathbf{k}^{(obs)}$ can account from the coarse scales, $\mathbf{t}_b^{(obs)}$ can “fill in” the information on the finer scales. This is seen in the posterior for w_{30} jointly conditioned on $(\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)})$. Bottom right: Note that the posterior density of $\ln(\delta)$ obtained from $(\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)})$ as well as just $\{\mathbf{t}_b^{(obs)}\}$ are skewed somewhat to the right of the prior, while that of the static-data-only inversion is skewed left. The prior and posteriors are truncated at $\ln(\delta) = 0, 4.6$ but the kernel density estimates used to create plots smooth them near the truncation limits. 45
- 8 Top left: Q-Q plot of the posterior of w_1 developed with both static and dynamic data (solid line) compared to that obtained from static data only (dotted line) and dynamic data only (dashed line). We see that there is little difference; the length-scale of the first Karhunen-Loève mode is large enough that the sensor grid provides a sufficiently complete sampling and breakthrough times contribute little. Top right: the posteriors for w_{15} , which show the impact of dynamic data. The inference drawn from static data only underpredicts the value of w_{15} . However, the inference drawn from just the dynamic data is not very informative. Bottom left: Inference of w_{30} is determined entirely by the breakthrough times since the length-scale of the Karhunen-Loève mode is too small to be sampled by the sensors in SSA. Consequently, the posterior developed from the static data only is indistinguishable from the prior. Strangely, the posterior conditioned on just the dynamic data is indistinguishable from the prior i.e., the dynamic data, by itself, does not inform very efficiently on the finer structures. Bottom right: The posterior density for $\ln(\delta)$ obtained from $\{\mathbf{t}_b^{(obs)}\}$ is indistinguishable from that obtained with $(\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)})$ 46

9	<p>Prior and marginalized posterior values of the Karhunen-Loève mode weights w_1, w_{15}, w_{30} and $\ln(\delta)$ as computed using the sensor-sets SSA (solid line) and SSB (dashed line). We see that the posteriors for w_1 are similar, indicating that the higher sensor density of SSB collects little extra information on w_1. However, the PDFs for w_{15} and w_{30} are quite different, indicating that the uncertainty in their values may have been underestimated in both cases. The posterior density for $\ln(\delta)$ obtained from SSB sensors is shallower than that obtained from SSA, though the maximum a posteriori estimate for $\ln(\delta)$ remains approximately unchanged.</p>	47
10	<p>Results from the posterior predictive check for breakthrough times conducted by generating $\mathbf{K}_{e,i}$ and using $\mathcal{M}(\mathbf{K})$ to generate the breakthrough times. Top: The filled, inverted triangles are the observations at the sensors. The median breakthrough times from the posterior predictive checks are plotted with open symbols and the error bars denote the 1st and 99th percentiles of the breakthrough time distribution. Left: We plot the results and the observations for SSA. Right: We plot the results for SSB. We see that the difference between the 1st and 99th percentiles is generally smaller for the SSB results, indicating a reduction in predictive uncertainty. Bottom: We plot the verification rank histogram for the ranks of the observations, given the predictive distribution of breakthrough times. Ideally, the histograms should be uniform; however, the histogram for the SSB predictions is marginally better (see the deviation-from-uniformity score in Table 1).</p>	48
11	<p>Fine-scale realizations of the random binary field conditioned on the observations and showing a single flowpath from the injection wells to each SSA sensor. These were developed from the posterior distribution of $\{\mathbf{w}, \delta\}$ using the mG-based technique described in Sec. 3.3. The center image (e) is the ground truth fine-scale realization. White indicates high permeability and gray is low permeability. Variation in the binary patterns are due to the stochastic nature of the process and the variation in the estimated δ parameter between realizations.</p>	50
12	<p>Comparison of the median travel times to the SSA sensors for three different data combinations. Left: Breakthrough time CDFs are shown for an example location (sensor 14). Right: The circle plots summarize the breakthrough time distributions and compare them to the true breakthrough times for all sensor locations. Details are in the text. “Coarse & fine” refer to realizations conditioned jointly on $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$, “Coarse only” refer to inferences using on $\{\mathbf{k}^{(obs)}\}$ while “Fine only” indicates inferences conditioned only on $\{\mathbf{t}_b^{(obs)}\}$.</p>	52
13	<p>Average absolute errors (AAE) between median travel times from the estimated and true travel time distributions for each field (left) and distribution of average KS statistic values calculated by comparing the full distribution of estimated and true travel time distributions (right).</p>	54

14	<p>Stages in the construction of the conditional PDF $P(\Delta\mathbf{K} \mathbf{F}, \delta)$. Top left: We create multiple realizations of the fine-scale random binary media field and plot their true upscaled log-permeability (as a cloud) and the model prediction $\mathbf{K}_e = \mathcal{L}(\mathbf{F}, \delta)$, for $\delta = 38$ grid-cells. Top right: we show the samples and the boxes from Level 4 of the KD-tree holding the boxed samples. Bottom left: We plot the kernel density estimate of $P(\Delta\mathbf{K}, \mathbf{F})$ for $\delta = 38$ grid-cells, constructed using Epanetchnikov kernels (after over-smoothing the bandwidths to remove isolated modes). Bottom right: We plot the conditional distribution $P(\Delta\mathbf{K} \mathbf{F})$, for $\delta = 38$ grid-cells, for various values of \mathbf{F}.</p>	59
15	<p>Q-Q plot for the posterior distribution of w_1 (left) and w_{30} (right). The solid line denote the Q-Q plot for inference developed with the KDE error model described in this section; the dashed line denote those that used the Gaussian model of Sec. 4. Only static data was used. The prior is plotted with dots. We note that the KDE model results in a plot that is not a straight line (for w_1); also the plots are different when the KDE and Gaussian models are used. Right, we see that the static data made no difference to the inference of w_{30}; the posterior is indistinguishable from the prior.</p>	62
16	<p>Posterior PDFs for $\{w_1, w_{15}, w_{30}, \ln(\delta)\}$ using the KDE error model, obtained with static data only (dotted lines), as well as static and dynamic data (solid line). We see that dynamic data affects the posterior for the higher Karhunen-Loève modes, as in Sec. 4. Also the impact of the static data is limited to the lower Karhunen-Loève modes. We see that the impact of dynamic data on $\ln(\delta)$ is low. We also plot the posterior PDF obtained using the Gaussian error model in Sec. 4. We see that at the finer (higher) Karhunen-Loève modes, which are impacted mostly by dynamic data, have posteriors that are relatively insensitive to the choice of KDE versus Gaussian data model.</p>	63
17	<p>Distributions of the tracer breakthrough times at the 20 SSA sensors, obtained via posterior predictive checks using the posterior density for $\{\mathbf{w}, \delta\}$. Those plotted in black were developed using the KDE error model; the ones in red use the Gaussian model described in Sec. 4. The observations are in black triangles; the circles are medians. The error bars indicate the 1st and 99th percentiles. We see that the width of the error bars are about the same irrespective of the error model (KDE/Gaussian) used; i.e. our choice of models did not reduce the uncertainty in the inference. However, only 2 observations fall outside the black error bars, whereas 3 do for the red ones, indicating a slight improvement in the goodness-of-fit when using the KDE error model.</p>	64

Tables

1	CRPS, MAE, IS and VRHD for the predictive ensembles of breakthrough times developed from the SSA and SSB sensor sets (see Fig. 10). CRPS, MAE and IS have units of breakthrough times i.e., they are non-dimensional. The VRHD was calculated as the mean of absolute deviations of the frequencies from one (the dashed line in Fig. 10). We see that the ensemble developed from the SSB set is somewhat better, reflecting the effect of a larger set of observations. The interquartile range was used to calculate IS.	39
2	Tabulation of the discrepancy between the observed and predicted breakthrough times at the 20 SSA sensors. All values are in dimensionless time. The predicted breakthrough time, in this context, is the median of the breakthrough times calculated over 1000 fine-scale realizations. This result corresponds to the offset of the circles from the center in Fig. 12. The “Nominal config.” is the configuration where the fluid is injected at the lower left corner of the domain and extracted at the upper right; the “Flipped config.” is the configuration where the injector is at the top left and the producer at the bottom right. $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ indicates fine-scale realizations developed from the inversion using both static and dynamic data; “ $\{\mathbf{k}^{(obs)}\}$ ” and “ $\{\mathbf{t}_b^{(obs)}\}$ ” indicate inferences drawn from just the static (“coarse-scale”) and dynamic (“fine-scale”) observations. Note that the observations were drawn only in the nominal configuration. We see that the posterior predictions are closer to observations in the $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ case, which use both the static and dynamic data, compared to the cases when only one type of data is used. Fine-scale realizations obtained from $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ data are also surprisingly accurate when predicting breakthrough times in the “flipped” configuration. This predictive robustness is missing in the realizations drawn from only one type of data – static (coarse-scale) or dynamic (fine-scale); the errors in the predictions for the “flipped” configuration are quite large.	55

- 3 Tabulation of the variation predicted breakthrough times at the 20 SSA sensors. Variation is defined as the width of the central 95% of the breakthrough times predicted by the 1,000 fine-scale realizations and corresponds to the radius of the circles in Fig. 12. The “Nominal config.” is the configuration where the fluid is injected at the lower left corner of the domain and extracted at the upper right; the “Flipped config.” is the configuration where the injector is at the top left and the producer at the bottom right. $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ indicates fine-scale realizations developed from the inversion using both static and dynamic data; “ $\{\mathbf{k}^{(obs)}\}$ ” and “ $\{\mathbf{t}_b^{(obs)}\}$ ” indicate inferences drawn from just the static (“coarse-scale”) and dynamic (“fine-scale”) observations. Note that the observations were drawn only in the nominal configuration. We see that the posterior predictions are closer to observations in the $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ case, which use both the static and dynamic data, compared to the cases when only one type of data is used. Fine-scale realizations obtained from $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ data are also surprisingly accurate when predicting breakthrough times in the “flipped” configuration. This predictive robustness is missing in the realizations drawn from only one type of data – static (coarse-scale) or dynamic (fine-scale); the errors in the predictions for the “flipped” configuration are quite large. . 56

Bayesian estimation of unresolved fine-scale properties of binary media from multiscale observations

1 Introduction

A binary medium is one that can be adequately described in terms of inclusions embedded in a matrix. The inclusions and the matrix can have vastly different properties, and the effective behavior of the binary medium arises as a nonlinear interaction of the dynamics of the binary medium’s two constituents. Further, the inclusions need not be distributed in the matrix evenly; also, their own properties, e.g., inclusion size, might vary in some structured manner over the entire domain. Thus, binary media can exhibit complex behaviors and rich dynamics. Examples of binary media are fractured rock (with the fractures acting as high permeability inclusions), sandstone layers with embedded shale, and engineered composite materials. The human skin, specifically, the stratum corneum, is another example; it consists of low permeability corneocytes separated by a lipid membrane (sometimes considered to be the “matrix”) which conducts organic compounds easily.

In many practical cases, e.g., fractured rock, the domain size of the binary medium (henceforth, the coarse-scale) may be a few orders of magnitude larger than the size of the inclusions. In such cases, it is impractical to measure the inclusions individually. However, the length-scale contrast between the inclusion and domain size is not large enough that the inclusions can simply be homogenized i.e., the domain cannot simply be represented by a mean value for the effect of the inclusions, but rather requires a more detailed characterization of the fine-scale. We call such behavior “multiscale”.

This multiscale behavior raises the possibility that it may be possible to infer the characteristics of the latent fine-scale from a judicious set of measurements conducted at the coarse and fine scales. In order to do so, one requires a *link function* \mathcal{L} , a model that locally captures the effect of fine-scale processes and structure at the coarse-scale. If this model is parameterized by structural/dynamical properties of the fine-scale, it becomes possible to construct statistical summaries for the fine-scale too. Once done, one may create realizations of the fine-scale that are equally consistent with observations at both scales; in the limit of an infinite number of error-free observations and with an error-free model, one should recover the true fine-scale.

We demonstrate inversion with a binary medium consisting of high-permeability inclusions in a low-permeability matrix. The permeabilities of the two materials, \mathcal{K}_i and \mathcal{K}_b , are known. We assume that we have noisy measurements of the effective log-permeability of the medium, $\mathbf{k}^{(obs)}$, at a few locations; they provide highly localized insights into the fine-scale structure and are referred to as *static data*. We also assume that we have noisy

measurements of breakthrough times, $\mathbf{t}_b^{(obs)}$, of a tracer at the same locations, from a pump test¹; they capture the integrated effect of fine-scale variations and are referred to as *dynamic data*. The proportion of inclusions $\mathbf{F}(\mathbf{x})$ are assumed to vary in space in an unknown but smooth manner with a known covariance structure. The inclusion sizes are assumed to vary randomly in space in an uncorrelated manner, and are represented by an unknown average size, δ . Together, $\mathbf{F}(\mathbf{x})$ and δ constitute a characterization of the fine-scale binary medium and their estimation from data $\mathbf{d} = \{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ constitutes a statistical summary of the fine-scale. Further, in conjunction with a model for generating fine scale structures given $\mathbf{F}(\mathbf{x})$ and δ , one can also develop realizations of the latent, unresolved fine-scale consistent with the observations.

The estimation of $\mathbf{F}(\mathbf{x})$ and δ is posed as a Bayesian inverse problem, predicated on a forward problem $\mathcal{M}(\mathbf{K})$ that predicts the breakthrough times \mathbf{t}_b at the observation locations. \mathbf{K}_e , the effective log-permeability field, is obtained from $\mathbf{F}(\mathbf{x})$ and δ via a link function i.e. $\mathbf{K}_e = \mathcal{L}(\mathbf{F}(\mathbf{x}), \delta)$. The link function \mathcal{L} is based on the upscaling of a random binary field created using excursion sets of multiGaussian fields convolved with a (Gaussian) kernel of size δ and truncated using $\mathbf{F}(\mathbf{x})$. To reduce the dimensionality of the inversion, we develop a reduced-order model (ROM) of $\mathbf{F}(\mathbf{x})$, based on a Karhunen-Loève (KL) decomposition of a 2D field. This allows us to generate $\mathbf{F}(\mathbf{x})$, and consequently $\mathbf{K}_e = \mathcal{L}(\mathbf{F}(\mathbf{x}), \delta)$, in a parametric manner by varying δ and $\mathbf{w} = \{w_i\}, i = 1 \dots M$ where M is the number of Karhunen-Loève modes retained in the ROM and w_i are their weights. Associated predictions of \mathbf{t}_b are obtained using $\mathcal{M}(\mathbf{K})$. The discrepancy between the observations (both static and dynamic) are attributed to errors, which are modeled either simply as i.i.d. Gaussian random variables (Sec. 4) or in a more sophisticated manner in case the errors are heteroscedastic (Sec. 6). The inverse problem is solved by sampling over the (\mathbf{w}, δ) space using an adaptive Markov Chain Monte Carlo (MCMC) technique and constructing a joint posterior probability density distribution $P(\mathbf{w}, \delta | \mathbf{d})$ from the samples. $P(\mathbf{w}, \delta | \mathbf{d})$ is thereafter used in posterior predictive checks, to construct fine-scale realizations of the binary medium, gather statistics on \mathbf{K}_e and \mathbf{t}_b at the observation locations and gauge the quality of the fit of the model to data.

The paper is structured as follows. In Sec. 2 we review literature on the key elements of the research presented here. In Sec. 3 we describe the forward model $\mathcal{M}(\mathbf{K})$, the link function \mathcal{L} , and models used for reducing the dimensionality of the inverse problem. In Sec. 4, we pose the inverse problem and test the inversion technique on a problem where homoscedastic errors are modeled as i.i.d Gaussian variables. In Sec. 6 we address heteroscedastic errors, caused by the approximate model for \mathbf{K}_e whose accuracy varies with $\mathbf{F}(\mathbf{x})$ and δ . We draw our conclusions in Sec. 7.

¹We define a pump test as follows: Water is transported through a porous medium by pumping under a steady pressure gradient. Once a velocity field has been established in the porous medium, a non-reactive tracer is injected in; the time it takes to arrive at various measurement points is called the *breakthrough* time. The exact configuration for the test that we will use in this paper is described in Sec. 4

2 Literature review

The estimation of field variables (permeabilities, hydraulic conductivities etc), per se, have long been topics of active research in hydrology; see [1, 2] for recent reviews. In this section we restrict ourselves to reviewing existing literature on the multiscale (or multi-level) inference of log-permeability fields modeled as random fields, methods for generating topologies for random binary fields, upscaling models for binary media and adaptive MCMC samplers.

2.1 Estimation of random fields

The use of random fields to regularize a spatial variable has been explored within the context of inferring log-permeability fields. Lee *et al.* [3] considered the estimation of spatially dependent permeability by modeling it as a random field. They discretized a rectangular domain with a Cartesian mesh, and estimated the permeability in each grid-block from dynamic data obtained from an “inverted 9 spot test”. Two separate prior models were adopted for the random field to regularize the problem – a Markov random field (MRF) model with an unknown precision parameter (i.e., the precision parameter was also inferred while estimating the permeability) and a model based on Gaussian processes (GP) with a known variogram and mean permeability in the rectangular domain. Neither of the two models reduced the dimensionality of the problem i.e., the number of parameters being estimated was equal to the size of the Cartesian mesh, which ranged between 32^2 and 64^2 . This required specialized updating schemes to improve mixing in the Metropolis-Hastings sampler used to construct the posterior distribution of the permeability field. About 50,000 samples were required. MRF priors were also used by Wang *et al.* [4] when estimating the initial (spatial) distribution of a contaminant embedded in a porous medium. The contaminant was transported by groundwater flows and time-variant concentration measurements were available at a few locations. They employed a hierarchical Bayesian formulation to estimate the concentration distribution as well as the precision of the MRF model and the variance of the measurement error. As in [3], no attempt was made to reduce the dimensionality of the inference problem. Fu and Gómez-Hernández [5, 6] present a more recent example of the use of MCMC with blocked-updating when inferring log-permeability fields. Unlike Lee *et al.* [3] where a red-black decomposition of grid-blocks was used to update the log-permeability field (modeled as a MRF) in the MCMC, they used a multi-Gaussian representation for the object of inference and devised a specialized technique for constructing the proposal within the MCMC. In particular, they divided the grid-blocks in the Cartesian mesh into concentric “strips”, which were updated together as a block in the MCMC; the proposals for the blocks were obtained by kriging the log-permeabilities in the strips surrounding the block in question. An independent sampler was used for the strips near the domain’s perimeter.

In field inversions, one often employs a parsimonious expansion of the object of inference in terms of a suitable set of bases. In [7], Li *et al.* address this problem by using a

Karhunen-Loève expansion of the covariance matrix of a random field. The authors consider the problem of estimating the log-conductivity of a reservoir from measurements of the steady-state hydraulic heads. The spatially variable log-conductivity field was modeled using a Gaussian Process (GP). The authors preserved 400 Karhunen-Loève terms in their expansion of the log-conductivity field and approximated the posterior distribution (i.e., conditional on head measurements) as a multivariate Gaussian. The posterior mean was estimated by posing a quadratic optimization problem and a lower bound on the posterior covariance was derived. They also performed a modified posterior predictive test due to Kitanidis [8] to gauge the quality of their inference. As part of their inference problem, the authors also estimate structural parameters – the variance and correlation lengths – but find that the posterior variance on the inferred parameters is large, in keeping with the observations of [8] and [3]. The use of Karhunen-Loève expansion to provide a reduced-order model for a random field was used by Marzouk *et al.* in [9] to estimate a log-diffusivity field. They considered a problem in 1D heat conduction in a domain with distributed (known) intermittent heat sources. They assumed that while the correlation length of the log-diffusivity field was known, the scale was not. The log-diffusivity field was modeled using 8 Karhunen-Loève bases, whose weights, along with covariance scale, were the objects of inference. The noisy measurements were that of temperature at a few “sensor” points; however, they were also functions of time. A model of nonlinear heat conduction was fitted to the measurements using a random-walk MCMC sampler. $O(10^5)$ MCMC sample were required to reconstruct the posterior distribution of the log-diffusivity field. The authors compared their method with a random-field representation that was not approximated i.e., the full covariance matrix was assumed known and used in the inference, resulting in an inverse problem of far higher dimensionality. As might be expected, the lower-dimensional inverse problem based on Karhunen-Loève expansions was easier and faster to solve; what was unexpected was that the Karhunen-Loève expansion, which was based on the *prior* covariance, proved to be such an efficient means of representing the *posterior* distribution of the log-diffusivity field.

The use of a basis set other than Karhunen-Loève modes has also been investigated. In [10], Jafarpur and McLaughlin compare the use of a Karhunen-Loève transform versus the discrete cosine transform (DCT) and find the latter to be more advantageous. In [11] they couple the reduced order DCT model to an ensemble Kalman filter to infer permeability fields as well as reservoir states via history matching. In [12], they use the discrete cosine bases as a *sparse* representation for the log-permeability field and infer their value *as well as the sparsity pattern* via history matching. A deterministic algorithm was used for fitting, while sparsity of the bases was ensured using a L_1 norm. To the authors’ knowledge, this is the first use of compressive sensing ideas in a nonlinear, inverse problem outside image and video processing.

Multiscale/multilevel inversion techniques seek to infer fine-scale structure of a field by assimilating data at multiple scales. This is because many types of measurements e.g., breakthrough times, often contain significant information regarding fine-scale structures of the medium being interrogated. However, performing such an inference, even after dimensionality reduction, using conventional techniques can be extremely challenging since

the forward problem has to be solved at a higher resolution to capture subtle fine-scale effects. Within the context of multilevel inference, the coarse-scale assimilation can be performed efficiently with low-resolution/low-fidelity models, and thereafter condition fine-scale models on inferences drawn at a coarser level. In [13] Ferreira *et al.* present a multilevel scheme for the inference of 1-D log-permeability fields. The log-permeability field was modeled as a multivariate Gaussian on each level. The primary novelty lay in the scheme for determining log-permeabilities at a level conditional on the log-permeabilities at an adjoining level. A somewhat simpler approach, though still based on a multilevel mesh, has been studied in [14]. The simplifications lie in the interlevel transfers – up-scaling is performed via a geometric average of permeabilities while downscaling involves kriging. Further, unlike in [13], only a single iteration between levels was performed. Three levels were used. A similar approach, but with two levels, was demonstrated in [15] where two separate Metropolis-coupled MCMC chains were run for the coarse and fine levels. This allowed the faster running (and better mixing) coarse model to scan and sample the log-permeability space efficiently; occasional swapping allowed the slower-running fine-scale MCMC chain to benefit from the coarse-scale chain’s properties. A similar approach with Metropolis-coupled chains was also investigated (with encouraging results) in [5] (“Scheme no. 5”).

Multiscale inference may also be performed by simply using a coarse/approximate model as a filter/preconditioner for MCMC proposals. Such an idea was investigated by Efendiev *et al.* [16] where they inferred the log-permeability field of a reservoir from production data. They used a Karhunen-Loève expansion of the log-permeability field (with 20 terms) to reduce the dimensionality of the inverse problem. They used a very coarse (low-resolution) model of flow through porous media to scan through the parameter space. Any proposed log-permeability field that was provisionally accepted by the coarse-scale model was re-run using the fine-scale model. The interpolation of the coarse log-permeability field to the finer one was simplified by their use of multiscale finite element bases in their forward problem. However, the authors mention that the acceptance rate was rather small, primarily due to the quality of the proposals in the MCMC chain. In [17] Efendiev *et al.* ameliorated the simulation by adopting a Langevin diffusion model so that their proposals would be tailored to the local topology of the posterior distribution. An added efficiency gain was identified by them due to the smooth nature of the pressure field that drove transport in the reservoir simulation. Velocities on the fine-scale were calculated using the “interpolated” coarse-mesh pressure, saving the expense of a large elliptic solve for pressure on the fine-scale mesh. The strategy of using a coarsened model to perform a preliminary scan of a proposal within the context of a MCMC scheme was also used in [5] (“Scheme no. 4”).

A combination of dimensionality reduction and multiscale inference has also been investigated by the use of “zonation” [18, 19, 20]. The procedure starts with dividing a domain into a small number of zones, identified using level sets and estimating hydraulic parameters by optimization, conditioned on observations. The parameters are assumed to be constant inside a zone. The level sets are further subdivided (in a manner similar to recursive mesh refinement) to improve the model fit; subdivisions that improve the fit are retained. The algorithm proceeds in an adaptive manner, with the model fit as the refinement met-

ric. Since zones can be of any size, small scale structures can, in principle, be inferred, as long as they improve the fit of the model to observations. However, the use of piecewise constant parameter values lead to non-physical jumps at zone boundaries. A version of the same, using smooth bases, has also been developed [21].

In our work, we borrow the approach adopted in [9, 16, 17] and use Karhunen-Loève expansions for the log-permeability field to reduce the dimension of the inverse problem. However, unlike [13, 15, 16, 17], we do not adopt a multilevel inversion scheme in an effort to resolve the fine-scales. Instead, we assume that a sufficient contrast exists between the resolution at which we perform the inference and the size of individual fine-scale structures (inclusions in our case) that a complete description of the fine-scale is not very useful. Consequently, we aim to infer *statistical summaries* of the fine-scale, conditional on the observations, and generate fine-scale realizations that could plausibly reproduce them. We do so by constructing a subgrid model to predict the effective log-permeability of a grid-block given the average inclusion proportion $\mathbf{F}(\mathbf{x})$ and size δ in a grid-block, as well as a model to generate realizations of the fine-scale random binary media given the same parameters. Models to develop fine-scale configurations are reviewed in Sec 2.2 and a particular model (which we will use in this work) is described in detail in Sec. 3.3. The problem of low acceptance rates of proposals in the MCMC sampler, which was solved by devising novel updating schemes in [3] or by Langevin random-walk inspired proposal distributions in [17], is addressed using a general, adaptive MCMC technique; this is described in Sec. 2.3.

2.2 Topological and upscaling models for random binary media

Models of flow through random binary media can be categorized approximately into (a) models that seek to estimate an effective permeability/conductivity of the media and (b) models that seek to generate geometries/realizations of such media. We discuss both below. Note that we will not address models for fractured media. Fractures are often modeled as discrete, thin but long high-conductivity entities (lines or plates) embedded in a low conductivity matrix or as fractured zones in continuum models. However, since fractures may span distances of the same size as the domain, they are at odds with our definition of inclusions being too small to resolve and will not be considered in this study.

Models for estimating permeability: There has been substantial work in estimating the effective permeability of a random, porous binary medium. Called effective medium theories (EMT) for binary assemblages [22, 23, 24, 25], these models seek to calculate a representative permeability as a function of the permeabilities of the two components and their relative abundances. Initial theories assumed spherical inclusions, though more recent studies have included more information on inclusion shapes [26, 27, 28, 29, 30, 31]. Reviews and comparisons to other methods can be found in [32, 33]. EMT-like approaches assume non-interaction between inclusions i.e., inclusions may not form long, interconnected flow-paths of high permeability, nor do they consider interaction between two components of the medium. Consequently, such models work best when the inclusion proportion is small e.g., less than 50% [34, 32].

Knudby *et al.* [33] have developed a model for estimating the effective permeability of a binary medium which *does* take the interaction of inclusions into consideration. It does so by using an average inter-inclusion distance as a proxy for the average length of a flowpath through the low conductivity component. Like the models described above, it also requires the relative proportions of the two components as input, along with aspect ratio and size of the inclusion. However, Knudby *et al.*'s model, taken as-is, requires creation of a realization of the binary medium so that the geometrical parameters (*viz.*, the inclusion size and the inter-inclusion distance) may be calculated. Within the context of an MCMC sampler, this can get prohibitively expensive and a significant adaptation is required before the concepts in [33] can be used to create a lightweight subgrid model for the permeability of binary media. These adaptations revolve around ways of estimating geometrical properties of inclusions and consequently, below, we review existing models for generating topologies/realizations of binary media.

Models for the geometry of random binary media: A very common means of generating representations of random media with multiple components is via indicator variables which choose between the components. Indicator geostatistical techniques [35] do so via a variogram to define spatial variations and can lead to very efficient simulation of binary media [36, 37, 38]. Alternatively, indicator simulation approaches may also be based on transition probabilities between indicator classes [39, 40]. Less common means of generating spatial binary media are object-based and Boolean models [41] and random media models based on pluriGaussian and truncated multiGaussian fields [41, 42].

The development of excursion set theory, as applied to multiGaussian fields, has mostly occurred in medical imaging and astrophysics [43, 44, 45]. In particular, given a *definition* of a multiGaussian field (note, *not* a realization), it is possible to analytically calculate expected values of the total area, the number of distinct excursions and the average excursion size above a given threshold [46, 47]. If multiple thresholds are specified, one may obtain a medium with inclusions of multiple components [48]. These ideas have also been employed in porous media flows - Phillips and Watson [49] proposed mean threshold crossing distances to estimate correlation lengths of permeability. However, the use of excursion sets of multiGaussian fields for characterizing and modeling heterogeneous media in groundwater flows is rare.

The review above raises the possibility of modeling a random binary medium as a truncated multiGaussian field and using excursion set theory to estimate expected values for inclusion size and number. Such a model has been proposed by the authors [50]. It predicts the effective permeability of a binary medium given the proportions of the two components, their individual permeabilities and a length-scale which is a proxy for its size. Being based on the expected values predicted by excursion set theory, the predictions are necessarily approximate since the inputs to the model can support an infinite number of random binary media configurations, each with its own permeability (though the scatter about the mean may not be much). Since the model is analytical, it is sufficiently lightweight for use in an inference scheme based on MCMC sampling. This model is described in more detail in Sec. 3.3

2.3 Adaptive Markov chain Monte Carlo techniques

In recent years, MCMC techniques have been increasingly used to fit models to observations [51], since they allow estimation of parameters while simultaneously quantifying the uncertainty in the estimate. Metropolis-Hastings (MH) samplers [51] are commonly used since they place no restrictions on the kind of models, the type of likelihood expressions and priors used in posing the inverse problem. Given starting values for model parameters, MH samplers “drift” the model parameters toward more likely values. Most simply, this is done by proposing parameter values, one at a time (called single-component update), from a Gaussian distribution centered at the current position in the parameter space and with a pre-specified proposal distribution. Blockwise updates, when a number of (or all) parameters are proposed all at once, by sampling from a multivariate Gaussian, are also widely used [3, 5, 51, 52, 53, 54].

As mentioned in [3] single-component updates lead to subpar mixing (exploration of the parameter space) of the MCMC chain; this is because variation of a single parameter scarcely makes any difference to the model predictions of a high-dimensional model. Blockwise updates often result in the other extreme, resulting in parameter combinations which are unacceptable. The most efficient proposal distributions are those that resemble the posterior [51], but this is generally the object of inference. In [3] separate blockwise updating schemes were devised for GP and MRF models of the log-permeability field. Specialized proposals (e.g., inspired by Langevin random walk) were devised in [17]. While our formulation uses the Karhunen-Loève coefficients \mathbf{w} and δ , rather than log-permeabilities, as the objects of inference, the same mixing and acceptance issues are expected to persist.

Adaptive Metropolis [55, 52] (AM) is a variation of the MH sampler which uses a *global* adaptive strategy to perform online tuning of the proposal to increase mixing and acceptance rates. The proposal is chosen from a multivariate Gaussian. AM starts with a pre-specified proposal density but periodically recalculates an empirical posterior covariance based on the samples collected up to that point. Typically, the pre-specified starting covariance is small so that samples are easily collected; as more samples are collected, the covariance asymptotically resembles that of the posterior. The technique is neither Markovian nor reversible and in [52] the authors identify the conditions under which AM will recover the desired stationary distribution. Delayed rejection [56, 57, 58] is a *local* MH variation that combines different proposals. One starts a MH sampler with a rather large proposal covariance. When a proposal is rejected, one simply scales down the initial covariance by a uniform factor, and tries again rather than simply advancing in the sample path (hence delayed rejection). The scale-down-and-retry can be performed as many times as needed, though generally a cut-off is enforced. DR has been shown to outperform MH [55]. Delayed Rejection Adaptive Metropolis (DRAM), the MCMC technique that we will use in this work, is an amalgamation of the AM and DR [53]. Sampling using DRAM progresses using DR, except the chain is periodically interrupted to recalculate an empirical covariance using the sample path. This allows a more accurate multivariate Gaussian to serve as the proposal distribution. Since DR is generally started with a rather large proposal covariance, the initial empirical covariance estimates (in AM) are expected to be “ambitious”,

where the scale-down-and-retry characteristic of DR can be profitably used to arrive at a reasonable estimate of the posterior covariance for use in the proposal distribution. DRAM is non-Markovian and proof of ergodicity i.e., it yields asymptotically unbiased estimators, can be found in [53].

This page intentionally left blank

3 Models used in the inverse problem

In this section we describe the models used in our inverse problem. We first review Karhunen-Loève expansions of random fields which are used to reduced the dimensionality of the inverse problem described in Sec. 4.1. Thereafter, we describe the models $\mathcal{M}(\mathbf{K})$ and \mathcal{L} which serve as the forward problem in the inversion.

3.1 Karhunen-Loève expansions of random fields

Let $R(\mathbf{x}, \omega)$ be a real-valued random field with zero mean, finite second moments and covariance function that is continuous in $\mathcal{D} \times \mathcal{D}$. $\omega \in \Omega$, where Ω is a sample space and $R(\mathbf{x}, \omega)$ can be considered to be a collection of real-valued random variables, indexed by $\mathbf{x} \in \mathcal{D}$, \mathcal{D} being a bounded spatial domain. Then, the Karhunen-Loève expansion of $R(\mathbf{x}, \omega)$ can be written as

$$R(\mathbf{x}, \omega) = \sum_{i=1}^{\infty} w_i(\omega) \sqrt{\lambda_i} \phi_i(\mathbf{x})$$

This equality holds in the pointwise and mean-square sense; convergence is in $L^2(\Omega)$ for all $\mathbf{x} \in \mathcal{D}$. Further, if $R(\cdot)$ is Gaussian and almost surely continuous, then the convergence is uniform in \mathcal{D} with probability 1 [59]. λ_i and $\phi_i(\mathbf{x})$ are the eigenvalues and eigenfunctions of the covariance kernel $C(\mathbf{x}, \mathbf{y})$

$$\int_{\mathcal{D}} C(\mathbf{x}_1, \mathbf{x}_2) \phi_i(\mathbf{x}_2) d\mathbf{x}_2 = \lambda_i \phi_i(\mathbf{x}_1). \quad (1)$$

Since $R(\cdot)$ is assumed Gaussian, the covariance kernel $C(\mathbf{x}, \mathbf{y})$ is symmetric and positive semi-definite and so, by [60],

$$C(\mathbf{x}_1, \mathbf{x}_2) = \sum_{i=1}^{\infty} \lambda_i \phi_i(\mathbf{x}_1) \phi_i(\mathbf{x}_2)$$

where $\phi_i(\mathbf{x})$ are continuous functions and form an orthonormal system in $L^2(\mathcal{D})$. Also, $w_i \sim \mathcal{N}(0, 1)$ and independent of each other.

Karhunen-Loève expansions are optimal in the sense that for all the possible orthonormal bases for $L^2(\mathcal{D})$, the $\{\phi_i(\mathbf{x})\}$ that satisfy Eq. 1 minimize the mean-squared error in a finite linear representation of $R(\cdot)$ [61] i.e. they minimize

$$\int_{\Omega \times \mathcal{D}} \left[R(\mathbf{x}, \omega) - \sum_{i=1}^M w_i(\omega) \sqrt{\lambda_i} \phi_i(\mathbf{x}) \right]^2 dP(\omega) d\mathbf{x} \quad (2)$$

for any $M > 1$.

In order to reduce the dimensionality of our inverse problem, we will model fields with a truncated Karhunen-Loève series. In such a case, the approximate field $R_M(\cdot)$ and its covariance function can be represented as

$$R_M(\mathbf{x}, \omega) = \sum_{i=1}^M w_i(\omega) \sqrt{\lambda_i} \phi_i(\mathbf{x}), \quad C_M(\mathbf{x}_1, \mathbf{x}_2) = \sum_{i=1}^M \lambda_i \phi_i(\mathbf{x}_1) \phi_i(\mathbf{x}_2)$$

The total variance or “energy” of $R_M(\cdot)$ is given by

$$\int_{\mathcal{D}} E(R_M(\mathbf{x}, \omega)^2) d\mathbf{x} = \int_{\mathcal{D}} C_M(\mathbf{x}, \mathbf{x}) d\mathbf{x} = \sum_{i=1}^M \lambda_i^2 \quad (3)$$

3.2 The transport model $\mathcal{M}(\mathbf{K})$

$\mathcal{M}(\mathbf{K})$ is a 2D Darcy-flow model for the transport of an inert tracer through an unsaturated porous medium by an incompressible, single-phase fluid. Given a log-permeability field \mathbf{K} , appropriate initial and boundary conditions (including a steady-state pressure gradient and a fluid source and sink inside \mathcal{D}), the model calculates a steady state velocity field and advects a tracer (treated as a passive scalar) through it to obtain breakthrough times \mathbf{t}_b at a set of N_s “sensor” locations inside \mathcal{D} .

As shown in Fig. 1(a), we consider a 2D domain \mathcal{D} with no-flow boundary conditions imposed on $\partial\mathcal{D}$. We consider a log-permeability field \mathbf{K} defined on \mathcal{D} . A fluid, with viscosity μ is pumped in at the lower left corner and pumped out at an equal rate at the upper right. In this problem, we will ignore the effect of gravity. Therefore, by Darcy’s law, the velocity \mathbf{v} is given by

$$\mathbf{v} = \frac{\mathcal{K}}{\mu} \nabla p,$$

where p is the pressure field defined on \mathcal{D} . The equation governing p (and therefore \mathbf{v}) is

$$\nabla \cdot \mathbf{v} = \nabla \cdot \frac{\mathcal{K}}{\mu} \nabla p = \frac{q}{\rho} \quad (4)$$

where q is the strength of the source/sink and ρ is the density of the fluid. The equation is solved using the second-order finite-volume scheme described in [62], on a uniform mesh. Two-point flux approximations are used, and the permeability at the interface of adjacent grid-blocks are estimated by a harmonic average. The solution yields p at grid-block centers and \mathbf{v} at the centers of grid-block edges i.e., the velocities are obtained on a staggered mesh. A solution of these equations, for a injection-production well pair in a binary medium, is shown in Fig. 1(b). The light areas denote scattered high-permeability inclusions in a dark, low-permeability matrix. A fluid is pumped in at the lower left corner and extracted at the upper right. The steady-state velocity field so achieved is used to plot a few streamlines that show the paths taken by the fluid in the binary medium to reach the exit. The medium is considered isotropic, and the permeability \mathcal{K} is modeled as a 2D field

rather than a full tensor. The proportion of high-permeability material $\mathbf{F}(\mathbf{x})$ and the true effective (upscaled) log-permeability are shown in Fig. 1 (c) and (d). Note that while $\mathbf{F}(\mathbf{x})$ shows a rather smooth variation in space, the \mathbf{K}_e field shows a rougher distribution and significantly more structure, indicating the significantly nonlinear nature of the mapping of $\mathbf{F}(\mathbf{x})$ to \mathbf{K}_e .

Solving Eq. 4 for p yields the velocity \mathbf{v} . This is used to advect an inert tracer using the model in Eq. 5.

$$\frac{\partial c}{\partial t} + \mathbf{v} \cdot \nabla c = q_c \quad (5)$$

where $c(\mathbf{x})$ is the concentration field of a tracer and q_c , defined only at the lower left and upper right corners of \mathcal{D} , is the source and sink for the tracer. The concentration of the tracer at the source, c_{source} , is set to 1 and $q_{c,source}$ is set equal to the flux of the fluid. The tracer flux at the sink, $q_{c,sink}$, is obtained by multiplying the fluid outflow with the local tracer concentration, $c_{sink}(t)$, which increases in time till it saturates at 1. Eq. 5 was solved on the same Cartesian mesh as Eq. 4, using an upwind second-order, finite-volume scheme [63]. Tracer concentrations were monitored at two sets of sensor locations (set A with 20 sensors and B with 34) as shown in Fig. 1(e) and (f); the time t_b at which $c = 0.5$ was achieved at any given sensor was denoted its breakthrough time. Note that Eq. 5 does not model pore-scale dispersion or molecular diffusion of the tracer.

3.3 The link function \mathcal{L}

In this section we summarize a recently developed statistical model (the link function $\mathcal{L}(\mathbf{F}(\mathbf{x}), \delta)$) to estimate the effect of unresolved inclusions on the log-permeability of a grid-block. Full details of this model can be found in [50]. Note that $(\mathbf{F}(\mathbf{x}), \delta)$ is an incomplete specification of the fine-scale structures that may reside in a grid-block and consequently an infinite number of realizations of the fine-scale may be conditioned to it. Each coarse-scale grid block has its own log-permeability, and the ensemble of log-permeabilities of the coarse blocks form a distribution. We hypothesize that one may analytically/semi-analytically derive a representative value for this distribution, which can then serve as a (deterministic) approximation for the ensemble; failing that, the model could be augmented with a simple, stochastic model for the discrepancy between model predictions and the “true” log-permeabilities of the realizations. Below, we develop a model for the representative value. It combines the upscaling method for binary media in [33] with concepts for representing random dispersed phases statistically.

Consider a binary medium composed of two components with permeabilities \mathcal{K}_l and \mathcal{K}_b . Consider, too, a rectangular inclusion of size $\Delta_x \times \Delta_y$, surrounded by a “box” of the binary medium (containing both the inclusion and the matrix) of size $l_x \times l_y$. Let \mathcal{K}_A and \mathcal{K}_H be the arithmetic and harmonic means of \mathcal{K}_l and \mathcal{K}_b . Then, per Knudby *et al.* [33],

$$\mathcal{K}_B^{-1} = \rho \mathcal{K}_A^{-1} + (1 - \rho) \mathcal{K}_H^{-1} \quad (6)$$

where

$$\rho = \frac{R - \mathbf{F}}{\frac{1}{\mathbf{F}} - \mathbf{F}} \quad \text{and} \quad R = \frac{\Delta_x l_y}{\Delta_y l_x}$$

Since inclusions are rarely rectangular, a more general means of calculating ρ has to be devised. Knudby *et al.* [33] proposed

$$\rho = \frac{2D_{norm} - D_{norm}^2}{1 - \mathbf{F}^2} \quad \text{where} \quad D_{norm} = \frac{l_x - \Delta_x}{\Delta_x}. \quad (7)$$

Thus D_{norm} denotes an average normalized distance traveled by the flow in the matrix between inclusions. For a particular realization of a random binary medium, a mean $\overline{D_{norm}}$ can be calculated by estimating a mean $\overline{\Delta_x}$ and $\overline{l_x}$. The distances between inclusions in the downstream direction of the flow can be measured and weighted by the sizes of the two inclusions to yield the mean $\overline{\Delta_x}$; similarly, a mean box length l_x can be calculated. However, the requirement of a particular realization of a random medium (i.e., resolving all the inclusions) makes a direct application of this method infeasible for inverse estimation.

We now develop a model for $\overline{D_{norm}}$ that does not require a random binary field, along with the inclusions, to be instantiated. Consider a spatially correlated multiGaussian (mG) field based on a Gaussian kernel

$$G(\mathbf{x}) = \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}d\Sigma^{-1}d^T\right) \quad (8)$$

where d is the position vector of any point in space measured from the kernel's origin. Such a field can be realized (e.g., on a mesh) by initializing an uncorrelated mG field (with values sampled i.i.d. from a standard normal) and convolving it with the Gaussian kernel in Eq. 8. We consider that the kernel is symmetric and aligned with the axes i.e., $\Sigma = \sigma^2 I$ is diagonal (I is the identity matrix). The symmetric Gaussian kernel used for the convolution is fully characterized by its full-width-at-half-maximum (FWHM), $\delta = 2\sqrt{2\ln 2}\sigma$. The FWHM also characterizes the spatial correlation of mG field and is a representative length-scale for the inclusions. If the correlated mG field is truncated at a threshold w , $-\infty \leq w \leq \infty$, we define a w -level excursion set as

$$X_w = \{x \in R^d : Y(x) \geq w\}.$$

The expression for the corresponding variogram can be found in [50].

Following [46], one can develop analytical expressions for the expected values of the following quantities - N , the number of cells above the truncation threshold w , the number m of distinct regions (inclusions) above the threshold and the number n of pixels in each

region. These are given by

$$\begin{aligned}
\frac{\mathbb{E}[N]}{S} &= \frac{1}{2\pi} \int_w^\infty \exp(-z^2) dz \\
\mathbb{E}[m] &= |EC| = \frac{\exp(w^2/2)}{(2\pi)^{3/2}} \left(\frac{\delta}{\sqrt{4\ln(2)}} \right)^{-2} |w| \\
\mathbb{E}[n] &= \frac{\mathbb{E}[N]}{|EC|}
\end{aligned} \tag{9}$$

Here, EC , the Euler Characteristic in 2D, represents the number of connected objects in the domain minus the number of holes in those objects. S is the area of the domain. The EC tends to 0 as w tends to 0, which in turn denotes $\mathbf{F} = 0.5$. For $w < 0$, $EC < 0$. A figure of the variation of EC with w can be found in [50].

The mean distance between the centroids of the inclusions is given by $\sqrt{(S/(\pi|EC|))}$. Since the inclusions have a non-zero area, the mean distance between inclusions is

$$\overline{D}^* = \sqrt{\frac{S}{\pi|EC|}} - 2\sqrt{\frac{\mathbb{E}[n]}{\pi}}$$

At $w = 0$ i.e., $\mathbf{F} = 0.5$, \overline{D}^* tends to δ , the expected size of the inclusions when they are equally abundant as the matrix. The modeled inter-inclusion distance is recast as

$$\overline{D}^* = \max \left(\sqrt{\frac{S}{\pi|EC|}} - 2\sqrt{\frac{\mathbb{E}[n]}{\pi}}, \delta \right) \tag{10}$$

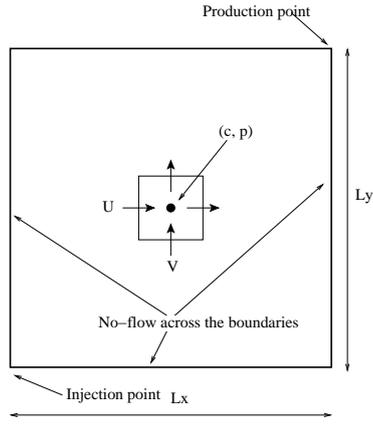
which is referred to as the basic model. Figures 2–4 in [50] show effective permeability \mathcal{K}_B modeled using the basic model for $\kappa = \log_{10}(\mathcal{K}_b/\mathcal{K}_l) = 2, 4$, for a number of values of δ and \mathbf{F} . They are compared to the permeability estimated numerically, using permeameter boundary conditions, for random binary media generated with those parameters. We see that (1) the largest errors occur around $\mathbf{F} = 0.5$, and the model overestimates the permeability for $\mathbf{F} > 0.5$. Larger δ also result in errors, especially as they become comparable to the size of the domain. Large values of κ e.g., $\kappa = 4$ also lead to inaccuracies. The basic model was adjusted for these parametric variations as:

$$\overline{D}^* = \begin{cases} \max \left(\sqrt{\frac{S}{\pi|EC|}} - 2\sqrt{\frac{\mathbb{E}[n]}{\pi}}, \frac{\delta}{\kappa-1} \right) & \text{if } \mathbf{F} \leq 0.5 \\ \max \left(\sqrt{\frac{S}{\pi|EC|}} - 2\sqrt{\frac{\mathbb{E}[n]}{\pi}}, \delta(1-\mathbf{F})^{\kappa-1} \right) & \text{otherwise} \end{cases} \tag{11}$$

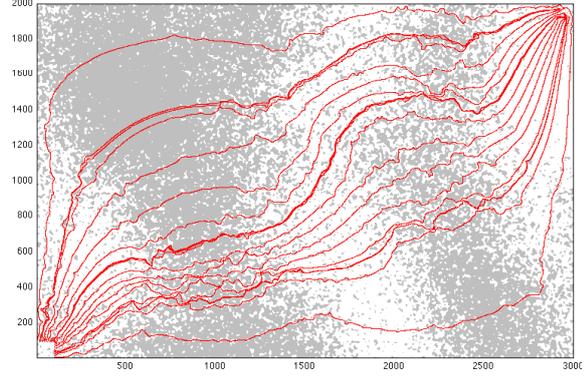
This allows us to model the effect of subgrid (unresolved) fine-scale structures on the permeability, as a function of \mathbf{F} and δ . Given $\mathbf{F} = \mathbb{E}[N]/S$, we evaluate w , $|EC|$ and $\mathbb{E}[n]$ using

Eq. 9. Thereafter, with knowledge of δ and κ , we estimate \overline{D}^* from Eq. 11, which is then used to calculate D_{norm} , ρ and permeability using Eq. 6. See [50] for comparison of the predicted permeability versus numerically upscaled ones obtained using the MODFLOW [64] finite-difference simulator.

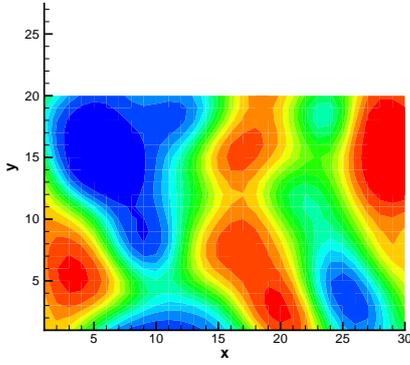
Note that this subgrid model is necessarily approximate. It predicts a log-permeability $\mathbf{K}_e = \mathcal{L}(\mathbf{F}, \delta, \kappa)$. The three parameters do not define a random binary medium uniquely and an infinite number of realizations are compatible with this specification. Thus the true permeability of random media (under)-specified in this manner will display a scatter about \mathbf{K}_e . In Fig. 2, we plot the effective log-permeabilities for $\mathcal{K}_l = 1$, $\mathcal{K}_b = 100$ and FWHM of 10 and 75 grid-cells as predicted by our model, \mathcal{L} . We also plot the numerically evaluated log-permeabilities corresponding to the same $\mathbf{F}(\mathbf{x}), \delta$ with points, forming a cloud around $\mathbf{K}_e = \mathcal{L}(\mathbf{F}(\mathbf{x}), \delta)$. The break in the log-permeability predictions $\mathcal{L}(\mathbf{F}(\mathbf{x}), \delta)$ at $\mathbf{F} \approx 0.5$ is due to percolation effects i.e., the inclusions start overlapping, creating high-permeability channels which are $O(l_x)$ in size. One may simply use \mathbf{K}_e as a deterministic approximation for the log-permeabilities of the various realizations of the binary field and model the discrepancy between observations of log-permeability and model predictions as simple i.i.d. Gaussians (homoscedastic errors). Alternatively, one may adopt a more sophisticated parameterization for the discrepancy/error (heteroscedastic errors). While the choice of the error model (Gaussian versus a more involved one) does not detract from the general characteristics of an inference procedure, it does impact the accuracy of the estimates/inferences. In the following sections, we explore the use of both simple and complicated error models to infer the permeability of a random binary medium, conditioned on limited observations.



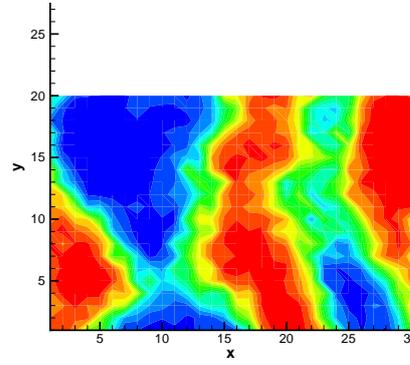
(a) Domain



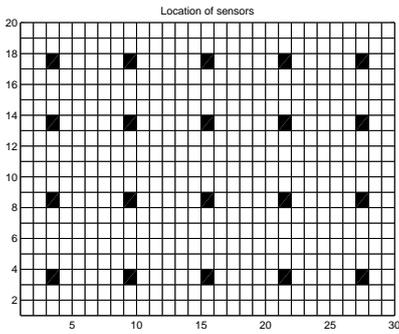
(b) Fine-scale



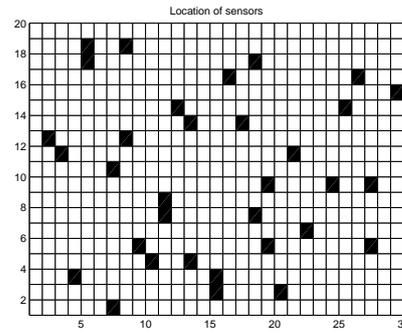
(c) Proportionality field $F(\mathbf{x})$



(d) The true $\log_{10}(\mathbf{K}_c^t)$ field



(e) SSA



(f) SSB

Figure 1: (a) The 2D rectangular domain \mathcal{D} with associated dimensions. No-flow boundary conditions are defined on $\partial\mathcal{D}$. A grid-block showing the collocation for pressure p and the velocities (u, v) in the x - and y -directions is shown. (b) We show a schematic of a binary medium with the light, higher permeability inclusions in a darker lower-permeability matrix. (c) The true $F_c^t(\mathbf{x})$ field for the proportion of inclusions. (d) The true upscaled log-permeability field \mathbf{K}_c^t . (e) The 30×20 coarse-scale computational mesh with the locations of 20 sensors (sensor-set A, SSA). (f) We show the locations of the 34 sensors in sensor-set B (SSB).

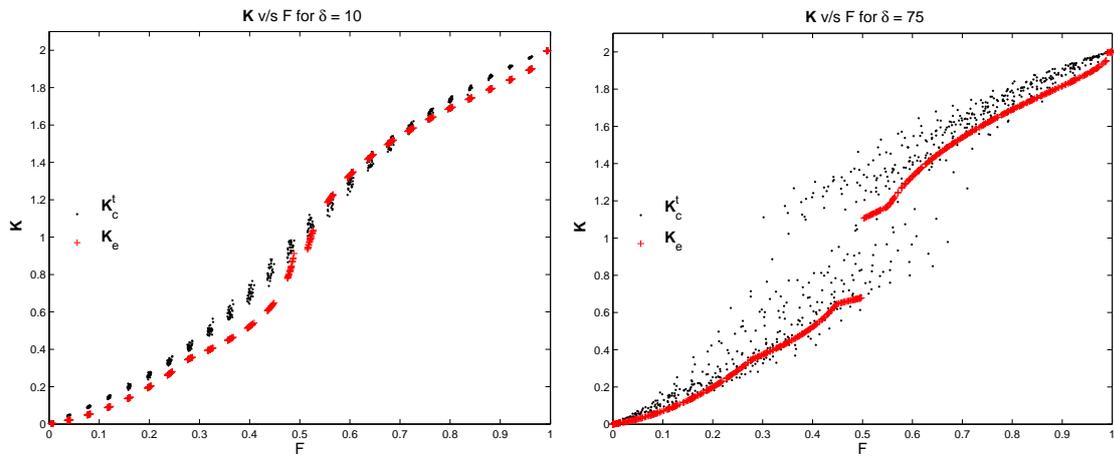


Figure 2: Plots of \mathbf{K}_e calculated using the link function \mathcal{L} , as well as numerically estimated from a random binary medium using MODFLOW-2005, as a function of $\mathbf{F}(\mathbf{x})$. Left: The plots are for a medium where $\delta = 10$ grid-cells. Right: $\delta = 75$ grid-cells. We see that the numerically estimated \mathbf{K}_e form a cloud around the value predicted by \mathcal{L} . This is because $\{\mathbf{F}(\mathbf{x}), \delta\}$ is an incomplete description of a random binary medium, and an infinite number of realizations, each with its own permeability, can be conditioned on such an incomplete specification. The break in \mathcal{L} predictions at $\mathbf{F} \approx 0.5$ is due to percolation effects.

4 Multiscale inference with homoscedastic errors

In this section, we pose and solve an inverse problem using the models described in Sec. 3. The thrust of this section is to demonstrate some general properties and capabilities of our approach. Specifically, we explore the ability of the observations to constrain the inference of structures at various spatial scales, the relative contributions of the static and dynamic data and the effect of increasing the observations. We will also demonstrate how the posterior distribution can be used to generate fine-scale realizations which are consistent with the observations.

We will model the discrepancy between true permeability and breakthrough times versus model predictions as i.i.d. Gaussian. This is somewhat simplistic and may not necessarily lead to accurate inferences. More sophisticated error modeling and accurate field reconstructions will be explored in Sec. 6.

4.1 Posing the inverse problem

Consider that the domain \mathcal{D} is 2D, rectangular, of size $L_x \times L_y$, and discretized by a Cartesian mesh of resolution $N_x \times N_y$. Consider a spatially variable field $\zeta(\mathbf{x})$, $-\infty \leq \zeta(\mathbf{x}) \leq \infty$ defined on \mathcal{D} . We model $\zeta(\mathbf{x})$ as a random field using GP, i.e., the discrete form of $\zeta(\mathbf{x})$, the vector ζ , is a random variable with multivariate Gaussian distribution and a known covariance matrix Γ . We further specify that the inclusion proportion $\mathbf{F}(\mathbf{x})$ is an analytical function of $\zeta(\mathbf{x})$, and it, in turn, governs the predicted effective log-permeability \mathbf{K}_e and breakthrough time \mathbf{t}_b via models $\mathcal{L}(\mathbf{F}(\mathbf{x}), \delta)$ and $\mathcal{M}(\mathbf{K})$. These are summarized below in Eq. 12

$$\begin{aligned}
 \zeta &\sim \mathbf{N}(0, \Gamma), \\
 \Gamma_{ij} &= C(\mathbf{x}_i, \mathbf{x}_j) = a \exp(-|\mathbf{x}_i - \mathbf{x}_j|^2 / b^2), \\
 \mathbf{F}(\mathbf{x}) &= \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{\zeta(\mathbf{x})}{\sqrt{2}} \right) \right), \\
 \mathbf{K}_e &= \mathcal{L}(\mathbf{F}(\mathbf{x}), \delta), \\
 \mathbf{t}_b &= \mathcal{M}(\mathbf{K}),
 \end{aligned} \tag{12}$$

where Γ_{ij} , an element of the covariance matrix Γ , denotes the correlation between grid-blocks i and j in the mesh. The scale a and range b of the covariance kernel are assumed known. The model $\mathcal{M}(\mathbf{K})$ is described in Sec. 3.2 and the link function \mathcal{L} in Sec. 3.3. Note that the analytical transformation $\zeta \mapsto \mathbf{F}(\mathbf{x})$ is required to map $\mathbf{F} \in [0, 1]$ to $\zeta \in [-\infty, \infty]$, so that a Gaussian model may be used. The inference does not depend on the particular transformation in Eq. 12; any transformation that allows the use of Gaussians (and consequently, a Karhunen-Loève decomposition) may be used.

Given a set of noisy data $\mathbf{d} = \{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ at a set of N_s ‘‘sensor’’ points, we wish to infer

the posterior distribution $P(\mathbf{K}_e, \mathbf{F}(\mathbf{x}), \delta | \mathbf{d})$. Using Bayes' formula, this can be written as

$$P(\mathbf{K}_e, \mathbf{F}(\mathbf{x}), \delta | \mathbf{d}) \propto P(\mathbf{d} | \mathbf{K}_e, \mathbf{F}(\mathbf{x}), \delta) \pi(\mathbf{K}_e, \mathbf{F}(\mathbf{x}), \delta) \quad (13)$$

where $\pi(\cdot)$ denotes our prior belief regarding the distribution of a variable. We model the discrepancy between the data \mathbf{d} and model predictions as Gaussians, i.e.,

$$\begin{aligned} \mathbf{k}^{(obs)} - \mathbf{K}_e &= \mathbf{k}^{(obs)} - \mathcal{L}(\mathbf{F}(\mathbf{x}), \delta) = \mathbf{e}_k \sim \mathcal{N}(\mu_k, \Gamma_k), \\ \mathbf{t}_b^{(obs)} - \mathbf{t}_b &= \mathbf{t}_b^{(obs)} - \mathcal{M}(\mathbf{K}) = \mathbf{e}_t \sim \mathcal{N}(\mu_t, \Gamma_t), \end{aligned} \quad (14)$$

where μ_k, μ_t are the means of the discrepancy between observed and model (i.e., predicted) log-permeabilities and breakthrough times and Γ_k, Γ_t the corresponding error covariances respectively. Under these (error) modeling assumptions, Eq. 13 reduces to

$$\begin{aligned} P(\mathbf{K}_e, \mathbf{F}(\mathbf{x}), \delta | \mathbf{d}) &\propto P(\mathbf{d} | \mathbf{K}_e, \mathbf{F}(\mathbf{x}), \delta) \pi(\mathbf{K}_e, \mathbf{F}(\mathbf{x}), \delta) \\ &\propto \exp\left(-[\mathbf{e}_k - \mu_k]^T \Gamma_k^{-1} [\mathbf{e}_k - \mu_k]\right) \exp\left(-[\mathbf{e}_t - \mu_t]^T \Gamma_t^{-1} [\mathbf{e}_t - \mu_t]\right) \\ &\quad \pi(\mathbf{K}_e, \mathbf{F}(\mathbf{x}), \delta), \end{aligned} \quad (15)$$

We assume that the discrepancies \mathbf{e}_k and \mathbf{e}_t are independent and model them as i.i.d. Gaussians with constant standard deviations i.e.,

$$\Gamma_k = \sigma_k^2 \mathbf{I}; \quad \Gamma_t = \sigma_t^2 \mathbf{I}.$$

where \mathbf{I} is the identity matrix. This is equivalent to stating that there are no systematic discrepancies between observations and model predictions as a function of $\mathbf{F}(\mathbf{x}), \delta$ and κ . Given the relationship between $\mathbf{K}_e, \mathbf{F}(\mathbf{x})$ and ζ (Eq. 12), and the modeling assumptions regarding errors, Eq. 15 can be compactly written solely in terms of ζ as

$$\begin{aligned} P(\zeta, \delta | \mathbf{d}) &\propto P(\mathbf{d} | \zeta, \delta) \pi(\zeta) \pi(\delta) \\ &\propto \exp\left(-\frac{[\mathbf{e}_k(\zeta) - \mu_k]^T [\mathbf{e}_k(\zeta) - \mu_k]}{\sigma_k^2}\right) \exp\left(-\frac{[\mathbf{e}_t(\zeta) - \mu_t]^T [\mathbf{e}_t(\zeta) - \mu_t]}{\sigma_t^2}\right) \\ &\quad \pi(\zeta) \pi(\delta). \end{aligned} \quad (16)$$

Here, we have also assumed that the prior distributions of δ and ζ are independent.

Solving the inverse problem in Eq. 16 would require us to infer each of the elements of ζ (though constrained by Γ); the dimensionality of the inverse problem is $N_x \times N_y$, which can be extremely large in case of fine meshes. In order to reduce the dimensionality of the inverse problem, we appeal to the discussion in Sec. 3.1 and construct a low-dimensional model of $\zeta(\mathbf{x})$ using a Karhunen-Loève expansion.

$$\zeta_M = \sum_i^M w_i \sqrt{\lambda_i} \phi_i, \quad \text{with } w_i \sim \mathcal{N}(0, 1), \quad (17)$$

where λ_i are the eigenvalues of the covariance matrix Γ , ϕ_i the corresponding eigenvectors, and w_i the weights which are modeled as i.i.d standard normals due to the GP model for ζ . In this expansion, we retain the M -largest eigenmodes. Note that by Eq. 3, the variance of the approximate field is less than the original one, and inferences will tend to be smoother than the true field. The link between the Karhunen-Loève weights \mathbf{w} and \mathbf{t}_b can be succinctly written as

$$\mathbf{w} \xrightarrow{\text{KL modes}} \zeta \xrightarrow{\text{Eq. 12}} \mathbf{F}(\mathbf{x}) \xrightarrow{\mathcal{L}, \delta} \mathbf{K}_e \xrightarrow{\mathcal{M}(\mathbf{K})} \mathbf{t}_b \quad (18)$$

Replacing ζ in Eq. 16 using its Karhunen-Loève expansion (Eq. 17), we get

$$P(\mathbf{w}, \delta | \mathbf{d}) \propto \exp \left(-\frac{[\mathbf{e}_k(\mathbf{w}) - \mu_k]^T [\mathbf{e}_k(\mathbf{w}) - \mu_k]}{\sigma_k^2} - \frac{[\mathbf{e}_t(\mathbf{w}) - \mu_t]^T [\mathbf{e}_t(\mathbf{w}) - \mu_t]}{\sigma_t^2} \right) \pi(\delta) \prod_{l=1}^M \exp(-w_l^2). \quad (19)$$

Here, the prior on \mathbf{w} , $\pi(\mathbf{w})$, has been expressed in terms of its independent elements, w_i , whose priors are standard normals. The objects of inference are the M elements of \mathbf{w} and δ . Note that the dimensionality of the inverse problem is no longer directly dependent on the mesh used for the inversion.

We will assume that the porosity of the binary medium is a constant in space and time, and is known. The basis for this assumption is the observation that variations in \mathbf{K} for most porous media are much larger than variations in porosity.

4.2 Solving the inverse problem

In this section, we solve Eq. 19 to develop realizations of both the fine- and coarse-scale fields that are consistent with the data. We first describe how we develop the “ground-truth” binary field (fine-scale), followed by the generation of the synthetic data that serve as observations, the development of the posterior distribution $P(\mathbf{w}, \delta | \mathbf{d})$ and finally, by an analysis of the inversion methodology.

Development of the fine-scale binary field: The domain \mathcal{D} is dimensioned as $L_x = 1.5, L_y = 1.0$. The covariance of the ζ field is initialized with $a = 1.0, b = 0.1(L_x^2 + L_y^2)^{1/2}$ (see Eq. 12). ζ is modeled as a multiGaussian field and a realization is obtained on a coarse 30×20 mesh of grid-blocks on \mathcal{D} . The “true” inclusion proportions on the coarse mesh, $\mathbf{F}'_c(\mathbf{x})$, is calculated using the transformation in Eq. 12. Each grid-block is thereafter further refined into 100×100 *grid-cells*, leading to a 3000×2000 grid-cell discretization of \mathcal{D} .

We generate a random binary field on the 3000×2000 mesh per the procedure in Sec. 3.3 and [50]. We initialize the mesh with a multiGaussian white noise and convolve it with a

Gaussian kernel with $\sigma = 5.0$ ($\delta = 11.774$) grid-cells. $\mathbf{F}(\mathbf{x})$, the proportion of inclusions in each grid-block (equivalent to $\mathbb{E}[N]/S$ in Eq. 9), is used to calculate the threshold w . The w -level excursion set creates the random inclusions in each grid-block. Repeated over all coarse-scale grid-blocks, we obtain the fine-scale binary medium on a 3000×2000 grid. The inclusions are assigned a log-permeability \mathcal{K}_b of 2 (i.e. a permeability of 10^2) while the matrix has a log-permeability \mathcal{K}_l of 0 (i.e., permeability of 1).

Generation of synthetic data: The synthetic data consists of measurements $\mathbf{k}^{(obs)}$ of effective log-permeability of the coarse grid-blocks containing the sensors in SSA and SSB. The permeabilities are calculated empirically by solving a permeameter boundary condition problem for each of the grid-blocks. For each grid-block, a time-independent pressure difference is imposed in one direction, and zero-outflow in the perpendicular one to calculate a flow-rate. This is used to calculate the effective grid-block permeability in that direction. The directions of pressure difference and zero-outflow are then exchanged to obtain the permeability in the perpendicular direction. The two permeabilities so calculated are similar and the harmonic mean of the permeabilities is taken as the “upscaled” permeability of the binary medium in that grid-block. Flow simulations for each of the grid-blocks in the 30×20 mesh are performed with MODFLOW-2005 [64]. We refer to the resulting log-permeability field as the true, upscaled log-permeability field, \mathbf{K}_e^t . Those values that correspond to the sensor grid-blocks in the SSA and SSB sets form the static data, $\mathbf{k}^{(obs)}$. I.i.d. Gaussian observation errors ($\sim N(0, 0.1)$) are added to them.

The original random binary field, on the 3000×2000 mesh is then subjected to a tracer transport simulation, as described in Sec. 3.2. MODPATH [65] is a Lagrangian particle tracking method operating on flux fields calculated in MODFLOW and is used here to simulate transport in the binary porous medium. A fluid is injected via a cross pattern of five wells within the coarse-scale grid-block in the lower-left corner and extracted via a similar configuration of pumping wells on the top-right grid-block to create a steady-state flow field. To calculate the breakthrough times, we reverse the steady-state velocity field. Then, for each coarse-scale sensor grid-block, we release 121 particles from a uniform 11×11 configuration of wells and advect those particles back to the injection wells using the reversed velocity field. The time it takes for half the particles (median of the distribution) released in a particular grid-block to reach the grid-block at the lower-left (the grid-block with the injection wells) is taken as the breakthrough time. Repeated for each of the sensor grid-blocks, we obtain the dynamic data $\mathbf{t}_b^{(obs)}$. I.i.d. Gaussian observation errors ($\sim N(0, 10^{-3})$) are added to them after non-dimensionalization (non dimensional time = (raw time)*(injection rate)/(domain pore volume)). Thus the breakthrough times contain the effect of all inclusions, fully resolved, on the dynamics.

The error model: The discrepancies between observations and model predictions are modeled as i.i.d. Gaussians. We assume that along with $\mathbf{k}^{(obs)}$, both \mathbf{F}^* and δ^* too can be measured at the grid-blocks with sensors, allowing us to estimate $\mathbf{K}_e^* = \mathcal{L}(\mathbf{F}^*, \delta^*)$ there. μ_k and σ_k are calculated empirically as the mean and standard deviation of $\mathbf{k}^{(obs)} - \mathbf{K}_e^*$ evaluated at the observation points. μ_t is set to zero. The standard deviation for the breakthrough time, σ_t , is set to 5% of the maximum non-dimensional breakthrough time observed in the

SSA set of sensors. These values are used in Eq. 16 and 19.

Generation of $P(\mathbf{w}, \delta | \mathbf{d})$: We solve Eq. 19 using the log form of δ , i.e. $\ln(\delta)$ which allows us to model its prior using a truncated Gaussian

$$\ln(\delta) \sim N(\ln(10), 2) \text{ if } 1 < \delta < 100.$$

Outside these limits, the prior is set to zero, thus loosely modeling the inclusions to be larger than a fine-scale grid-cell, but smaller than a (coarse) grid-block. 30 terms were retained in the Karhunen-Loève expansion of spatial field i.e., $M = 30$ in Eq. 17. 10^6 samples were taken using the adaptive MCMC sampler (DRAM, [53]) and the chain was checked for mixing and burn-in using the autocorrelation-based metric in [66] (as implemented in the `mcgibbsit` package [67] in R [68]) by monitoring the 5th and 95th percentiles as well as the median. 10,000 samples were retained by thinning the chain and used for developing posterior distributions of the objects of inference.

In Fig. 3 we plot the probability density functions (PDF) for w_1, w_{15}, w_{30} and $\ln(\delta)$, by marginalizing over the thinned samples. The three weights, w_1, w_{15}, w_{30} , correspond to 3 Karhunen-Loève modes which are representative of large, medium and small-scale (but resolved) structures. Their joint PDFs are also plotted. We see that the posterior distribution for w_1, w_{15} , and w_{30} are roughly Gaussian, though that does not hold true for $\ln(\delta)$. The median value of δ is found to be 10.25 grid-cells (compared to the true value of 11.774) with a 90% credibility interval of (1.4–72.24) grid-cells. The inter-quartile range is (3.76–27.3) grid-cells. The approximately Gaussian (posterior) distributions for w_{15} and w_{30} are centered around 1, unlike the priors which are standard normals. The scatter plots show that the samples of these three parameters are not correlated with each other; the correlations between the (posterior of the) weights of adjacent Karhunen-Loève modes weaken very quickly and are insignificant for modes which are 5 modes apart.

4.3 Results and discussion

In this section, we analyze the solution of the inverse problem. We first perform the inversion using both the static and dynamic data, and then repeat using the static and dynamic data individually. This is done to explore the contribution of each type of data to the inference.

4.3.1 Assessment of $\mathbf{F}(\mathbf{x})$ and \mathbf{K}_e inferences

We use the samples from $P(\mathbf{w}, \delta | \mathbf{d})$ (Eq. 19) to generate realizations of $\mathbf{F}_i(\mathbf{x})$ and $\mathbf{K}_{e,i}, i = 1 \dots 10,000$. The first 500 samples are discarded as “burn-in”. In Fig. 4, in the top row, we plot their expected values $\bar{\mathbf{F}} = \mathbb{E}[\mathbf{F}(\mathbf{x})]$ and $\bar{\mathbf{K}}_e = \mathbb{E}[\mathbf{K}_e]$ in color while the “true” values (as plotted in Fig. 1, middle row) are plotted as dotted contours. For $\bar{\mathbf{F}}$, we see that the shade plot and the contours are similar with respect to large scale structures, though there

are differences at smaller scales. On the other hand, the $\overline{\mathbf{K}}_e$ plot at the top right shows significant differences between the shade and contour plots. This arises mostly because of the sharp gradients in \mathbf{K}_e and the approximations inherent in the $\mathcal{L}(\mathbf{F}(\mathbf{x}), \delta)$ model for the log-permeability. In the middle row of Fig. 4, we plot the grid-block-wise standard deviation of the 9,500 samples of $\mathbf{F}_i(\mathbf{x})$ and $\mathbf{K}_{e,i}$, along with the locations of the sensors in SSA. We see that the standard deviations (the uncertainty in the inferences) are minimum at the measurement points and increase in the poorly instrumented regions. The regions with the largest uncertainties are concentrated in regions of high gradients where the smooth Gaussian-process based representation forms a poor model for the spatial variation. In the bottom row, we plot the difference between the true and average values i.e. $\varepsilon_F = \mathbf{F}_c^t(\mathbf{x}) - \overline{\mathbf{F}}$ and $\varepsilon_K = \mathbf{K}_c^t - \overline{\mathbf{K}}_e$. We see that sensor locations are generally regions of low error. In the plot for ε_K , we see the largest errors correspond strongly with the “wall” of high gradients in \mathbf{K}_e at $x = 25$, where the error approaches 50%.

In Fig. 5 we create a counterpart of Fig. 4 but using static data only. All results are drawn from 9,500 samples, identical to Fig. 4. In the top row, we plot the expected values $\overline{\mathbf{F}} = \mathbb{E}[\mathbf{F}(\mathbf{x})]$ and $\overline{\mathbf{K}}_e = \mathbb{E}[\mathbf{K}_e]$ in color while the “true” values (as plotted in Fig. 1, middle row) are plotted as dotted contours. We see that the inferences drawn from static-data only capture the large scale structures and miss the finer details. This results in smaller errors in $\overline{\mathbf{F}}$ which tends to be smoother than $\overline{\mathbf{K}}_e$. In the middle row, we plot the standard deviation of the 9,500 samples of $\mathbf{F}_i(\mathbf{x})$ and $\mathbf{K}_{e,i}$, along with the locations of the sensors in SSA. As expected, they are smallest at the sensor locations. Further, compared to their counterparts in Fig. 4, they are larger. In the last row, we plot the difference between the true and average values i.e. $\varepsilon_F = \mathbf{F}_c^t(\mathbf{x}) - \overline{\mathbf{F}}$ and $\varepsilon_K = \mathbf{K}_c^t - \overline{\mathbf{K}}_e$. Compared to their counterparts in Fig. 5, the errors are far larger. This arises from the inability of the inferences, drawn solely from static data (which are informative only on the larger scales) to reconstruct finer details. This is particularly true for the $\overline{\mathbf{K}}_e$ field, which shows regions of high gradient.

In Fig. 6, we plot the counterparts to Fig. 4 and Fig. 5 but using estimates drawn only from the dynamic data, i.e., the breakthrough times. We see that the inferences drawn are extremely poor and bear little resemblance to the true $\mathbf{F}(\mathbf{x})$ or \mathbf{K}_e fields which are plotted as contours over them. The standard deviations (middle row) and errors (last row) are correspondingly large, compared to their counterparts in Fig. 4 and Fig. 5.

4.3.2 Assessment of $\{\mathbf{w}, \delta\}$ inferences

Next we analyze the individual weights of the Karhunen-Loève modes as drawn from inferences conditioned only on static or dynamic data as well as the inferences that are conditioned jointly on both types of data. In Fig. 7, we plot the marginalized PDFs for w_1, w_{15}, w_{30} and $\ln(\delta)$ conditioned jointly on both types of data (static & dynamic), as well as those conditioned on just the static and dynamic data individually. We also plot the priors (using symbols) for comparison. The solid line refer to posteriors which were obtained using both the static and dynamic data; the dotted lines refer to inferences from the static data only and the dashed line is used to indicate estimates drawn from only the dynamic

data. We see, in the top left figure, that the observations are most informative about w_1 (i.e., the difference between the prior and posterior is the largest), the Karhunen-Loève mode corresponding to the largest structures. In fact, the grid spacing of the sensors in SSA is smaller than the length-scale of the first Karhunen-Loève mode and provides a very good sampling of it. Consequently, the distribution of w_1 obtained using just the static data is almost identical to that obtained from using both the static and dynamic data. Relative to the static data, the dynamic data contributes very little to the inference of the lower (larger) Karhunen-Loève modes - the posterior for w_1 in Fig 7 (top left) is barely different from the prior. For intermediate Karhunen-Loève models e.g., w_{15} , (top right in Fig. 7) both the static and dynamic data contribute to the inference, though the contribution of the static is much more; the posterior developed from just the static data is quite close that developed using $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$. At the small scales i.e., w_{30} (Fig. 7, bottom left), the dynamic data does not contribute much directly - the posterior distributions obtained from the $\mathbf{k}^{(obs)}$ -only and $\mathbf{t}_b^{(obs)}$ -only inversions are almost the same as the prior, whereas the inversion conditioned jointly on $(\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)})$ is informative. Thus dynamic data is informative, *only after* the larger/coarser scales have been accounted for by the static data; by itself, $\mathbf{t}_b^{(obs)}$ does not have the information content to resolve both. Transport occurs preferentially through high permeability regions, which can be affected by the smallest scales; since breakthrough times are the integrated effect of the travel times of the tracer, one may naively expect that the effect of small scale variations are easily captured there. However, breakthrough times are also affected by the larger Karhunen-Loève modes and deconvolving the impact of the fine and coarser scales cannot be done without $\mathbf{k}^{(obs)}$, which uniquely captures the coarse-scale structures. The inference with the SSA sensors collect some information about δ but there is a clear trend away from values of $\ln(\delta) < 2$ in inversions done using $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ as well just $\{\mathbf{t}_b^{(obs)}\}$ (Fig. 7, bottom right).

The marginalized posteriors plotted in Fig. 7 have a shape very similar to a Gaussian. Since the prior for $\{\mathbf{w}, \ln(\delta)\}$ are Gaussians (and those for \mathbf{w} , standard normals), quantile-quantile (Q-Q) plots are a convenient means for identifying the degree by which the posteriors deviate from a Gaussian distribution. Fig. 8 contains the Q-Q plots corresponding to distributions of Karhunen-Loève weights plotted in Fig. 7. The Q-Q plot for the prior are shown with symbols (a straight line with a slope of 1). We note that the plots are approximately straight lines, indicating that they do not deviate much from a Gaussian distribution. This is a surprising finding, given that the transport model is a nonlinear one, as is the link function.

4.3.3 Impact of the number of sensors

In Fig. 9 we compare the marginalized posteriors for w_1, w_{15}, w_{30} and $\ln(\delta)$ as obtained with sensor-sets SSA and SSB. Both coarse and fine-scale data are recorded for each sensor set. The second set, which has more than half as many sensors, results in a PDF for w_1 which is much sharper than that obtained with SSA, while the PDF for $\ln(\delta)$ is shallower

and wider than the one observed for SSA, with a maximum a posteriori estimate remaining unchanged. However, the PDFs for w_{15} and w_{30} change significantly when recomputed using SSB, indicating that the information content of $\mathbf{t}_b^{(obs)}$ collected by SSA and SSB could be different. There is also a possibility that the uncertainty in the inferences may have been underestimated when computed using SSA. The reason for this can be found in Fig. 2. We see that a given $\{\mathbf{F}(\mathbf{x}), \delta\}$ may support a distribution of \mathbf{K}_e ; expanded to a 2D field, a $\{\mathbf{w}, \delta\}$ characterization may support a range of \mathbf{K}_e fields and consequently \mathbf{t}_b at the sensors. In our inference scheme, the range of \mathbf{K}_e corresponding to a $\{\mathbf{w}, \delta\}$ characterization is not explored; rather we impute a permeability field $\mathbf{K}_e = \mathcal{L}(\mathbf{F}(\mathbf{x}), \delta)$. This deterministic mapping excludes regions in the $\{\mathbf{w}, \delta\}$ -space, which under \mathbf{K}_e -exploration in the MCMC could have resulted in $\mathbf{t}_b = \mathcal{M}(\mathbf{K})$ similar to $\mathbf{t}_b^{(obs)}$ and thus been accepted as valid samples for a (rather wide) posterior distribution.

In order to analyze the model fits arising from SSA and SSB measurements, we conduct a posterior predictive check (PPC) using $\mathbf{t}_b^{(obs)}$. Note that since inferences were drawn using both $\mathbf{k}^{(obs)}$ and $\mathbf{t}_b^{(obs)}$, good fits with $\mathbf{k}^{(obs)}$ may compensate for bad fits with $\mathbf{t}_b^{(obs)}$. In Fig. 10, we plot the results of PPC performed using inferences drawn from both SSA and SSB. We use the realizations $\mathbf{F}_i(\mathbf{x}), i = 500 \dots 10,000$, to generate the corresponding $\mathbf{K}_{e,i} = \mathcal{L}(\mathbf{F}_i(\mathbf{x}), \delta_i)$; thereafter $\mathbf{t}_{b,i} = \mathcal{M}(\mathbf{K}_{e,i})$ at the SSA- and SSB-sets of sensors. We plot the median, the 1st and the 99th percentile of the breakthrough times. We see that the spread of the replicates of \mathbf{t}_b obtained with SSB (right figure) is generally smaller than those obtained with SSA, indicating a reduction in predictive uncertainty.

We next evaluate the predictive ensemble of breakthrough times (plotted for SSA and SSB in Fig. 10) quantitatively using the continuous rank probability score (CRPS), the mean absolute error (MAE) and the interval score (IS) [69, 70]. 950 samples from the predictive ensemble (rather than the full 9500 ensemble members) were used for the purpose; the sensitivity of the metrics to the number of samples was checked by repeating the calculations with double and half the number of samples. The interquartile range was used for calculating the IS. In Fig. 10 (bottom row) we plot the verification rank histogram (VRH) for the observations, in the breakthrough time predictive ensembles developed from the SSA and SSB sets of sensors. Each observation, along with the predicted breakthrough times, is sorted and ranked; the ranks of the 20 (34, in case of SSB) observations are histogrammed into 20 (34) bins. Ideally, the VRH should indicate a uniform distribution, with a frequency of one for each bin; deviations from uniformity (VRHD) are a measure of lack of calibration. This can be due to lack of data as well as shortcomings of the model itself. We see that the predictive ensembles developed from both SSA and SSB sets are somewhat under-dispersive (the observations are ranked at the extremities). In Table 1, we tabulate the CRPS, MAE, VRHD and IS for the two predictive ensembles in Fig. 10; the SSB ensemble is slightly better. Thus the extra observations in the SSB set lead to a more accurate estimate of $\{\mathbf{w}, \delta\}$, which, in turn, results in a more predictive ensemble, as plotted in Fig. 10 (right). Note that a naive check of the posterior distribution would have led one to believe that SSB provides a *worse* model fit — while the PPC results for SSA show that 3 observations fall outside the 98% credibility interval, in SSB the corresponding number

is 6. This, however, is compensated for by the tighter distributions and the fact the SSB has almost twice as many observations as SSA.

Table 1: CRPS, MAE, IS and VRHD for the predictive ensembles of breakthrough times developed from the SSA and SSB sensor sets (see Fig. 10). CRPS, MAE and IS have units of breakthrough times i.e., they are non-dimensional. The VRHD was calculated as the mean of absolute deviations of the frequencies from one (the dashed line in Fig. 10). We see that the ensemble developed from the SSB set is somewhat better, reflecting the effect of a larger set of observations. The interquartile range was used to calculate IS.

Ensemble	CRPS	MAE	IS	VRHD
SSA	0.047	0.0625	0.239	1.1
SSB	0.046	0.0593	0.186	0.88

4.4 Summary

To summarize, we have developed an inference scheme, predicated on a multiscale link function, to infer coarse-scale features and summaries of fine-scale structures of a random binary field from a combination of static and dynamic observations. We can estimate a spatially variable inclusion proportion $\mathbf{F}(\mathbf{x})$ as well as a representative value for the inclusion size δ from observations that are obtained on the coarse-scale but are nevertheless informative about the fine-scale. We see that the static data can inform on the large-scale features mainly because the distribution of sensors is sufficient to resolve such structures. The dynamic data, which is an integrated measure of the effect of small variations in the permeability field is key to estimating smaller (but nevertheless resolved) structures. Information on subgrid variations is obtained by the link function \mathcal{L} ; its parameters provide summaries of the unresolved scales. Increasing the number of sensors leads to estimates with lower predictive uncertainties. However, posterior predictive checks of our model reveal that the fit to data could be improved; the quality of fit is mostly determined by the quality of our multiscale link model \mathcal{L} and the crude i.i.d. Gaussian error model for the discrepancy between observations of log-permeability and model predictions, though increasing the number of observations does improve the inferences.

We have also explored how the static ($\mathbf{k}^{(obs)}$) and dynamic ($\mathbf{t}_b^{(obs)}$) data contribute to the estimation of $\{\mathbf{w}, \ln(\delta)\}$. By itself, $\mathbf{t}_b^{(obs)}$ is not sufficiently informative to estimate both the large-scale structures and the finer details. However, if the large-scale structures are constrained/estimated using $\mathbf{k}^{(obs)}$, the finer details can be inferred from $\mathbf{t}_b^{(obs)}$. Thus joint inversions on $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ result in more accurate estimations of $\{\mathbf{w}, \ln(\delta)\}$ not only because of the larger number of observations involved, but rather because of the *type* of information. Since $\mathbf{k}^{(obs)}$ is informative about the larger lengthscales in the domain while $\mathbf{t}_b^{(obs)}$

is impacted most by the finer scales that contort the flowpaths in the porous medium, joint inversion conditions the estimates to *multiscale* data.

In the next section (Sec. 5), we will explore the robustness of the inference. Since the aim of reconstructing a permeability field is generally to use it to predict transport phenomena, we will subject the reconstructions conditioned on multiscale data (as well as those obtained individually from static and dynamic data) to posterior predictive tests. We will show that multiscale inversions result in permeability fields that reproduce observations far more accurately compared to the other two cases. Further, these permeability fields retain their enhanced accuracy in flow conditions which are different from the injector/flow configuration which were used to obtain $\mathbf{t}_b^{(obs)}$ which were used to estimate the permeability field. We also show that this does not hold true for the other two cases.

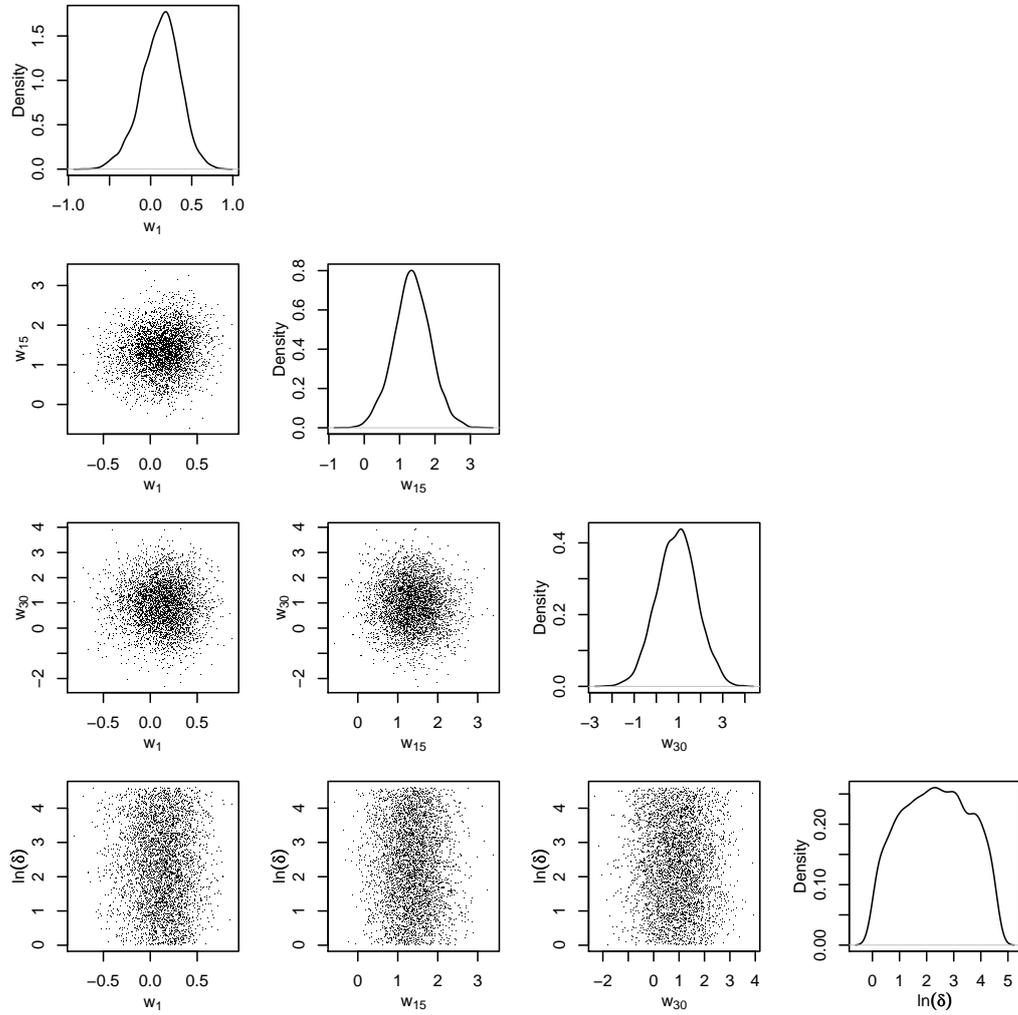


Figure 3: Plot of the joint and marginal probability densities of w_1, w_{15}, w_{30} and $\ln(\delta)$ from the posterior distribution. The joint distributions show very little correlations between the Karhunen-Loève modes at the large, medium and small scales, as well as their correlations with δ . Also, PDF of the individual parameters are roughly Gaussian, with the exception of δ .

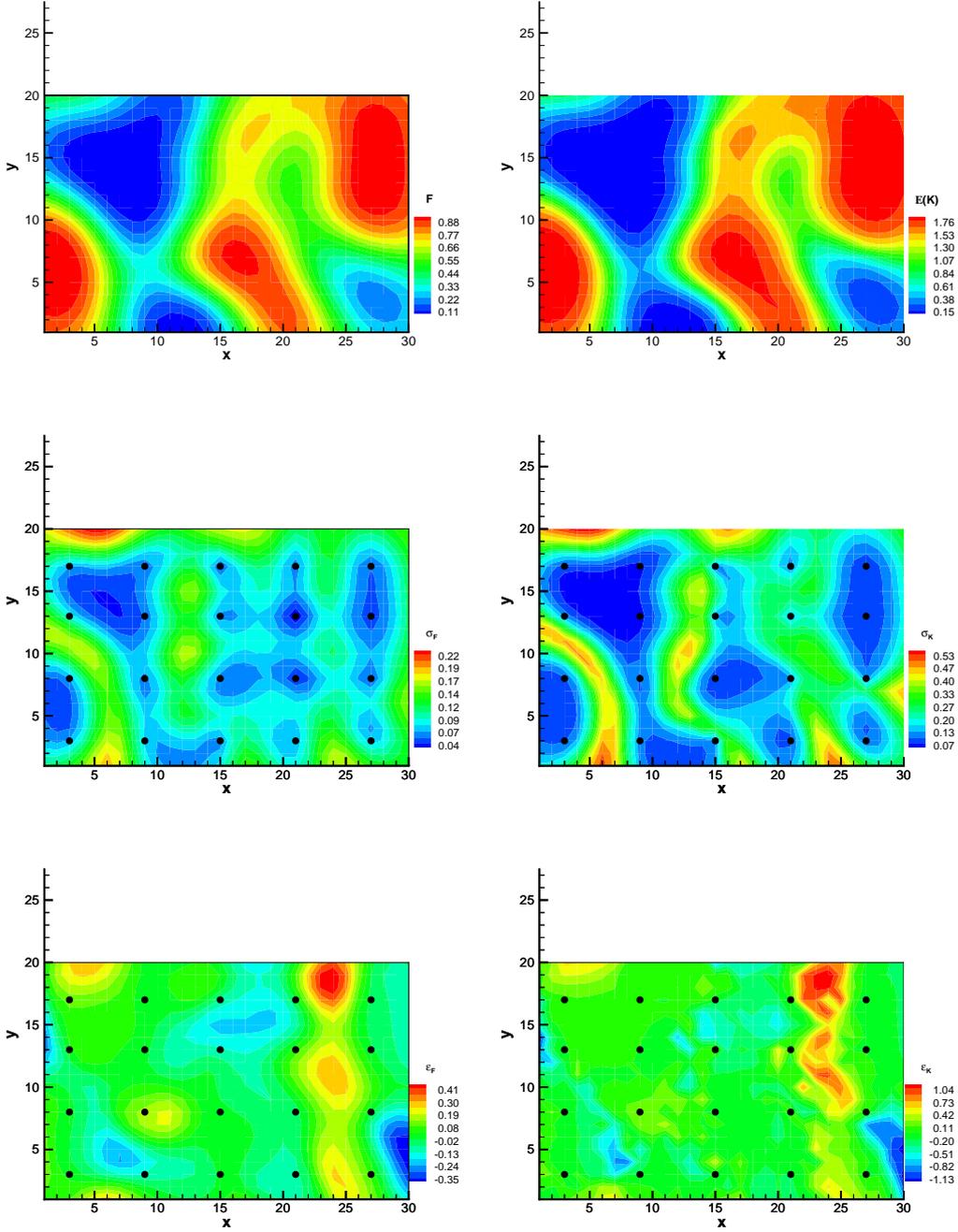


Figure 4: Results from the inference, computed using 9500 samples from the posterior distribution. In the top row, we plot the means $\bar{\mathbf{F}} = \mathbb{E}[\mathbf{F}(\mathbf{x})]$ and $\bar{\mathbf{K}}_e = \mathbb{E}[\mathbf{K}_e]$; they bear a strong resemblance to the true values in Fig. 1 (middle row). In the middle row, we plot the grid-block-wise standard deviation of $\mathbf{F}_i(\mathbf{x})$ and $\mathbf{K}_{e,i}$; they are smallest at the sensor locations. In the bottom row, we plot the errors $\epsilon_F = \mathbf{F}_c^t(\mathbf{x}) - \bar{\mathbf{F}}$ and $\epsilon_K = \mathbf{K}_c^t - \bar{\mathbf{K}}_e$, which show large values in regions of high gradient where the representation errors due to the smooth Gaussian process model are the largest.

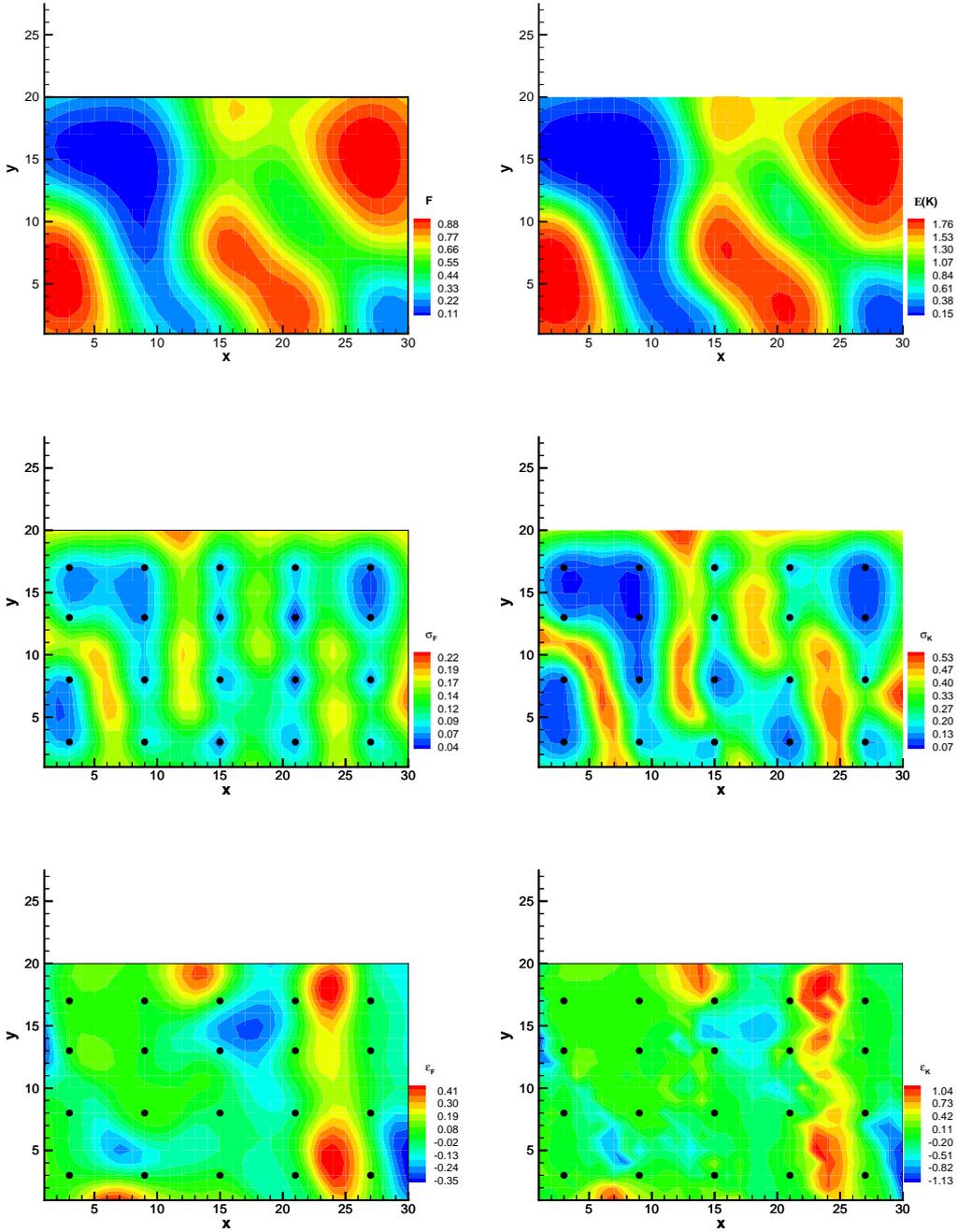


Figure 5: Results from the inference, computed using 9500 samples from the posterior distribution. Only static data was used to draw the inferences. In the top row, we plot the means \bar{F} and \bar{K}_e which are worse than those seen in Fig. 4. Further, \bar{K}_e has larger errors than \bar{F} due to its finer structures. In the middle row, we plot the grid-block-wise standard deviation of $F_i(\mathbf{x})$ and $K_{e,i}$; they are larger and more widely distributed than their counterparts in Fig. 4. In the bottom row, we plot the errors ϵ_F and ϵ_K , which show larger values in regions of high gradient. Further, the errors are higher and more widely distributed compared to their counterparts in Fig. 4.

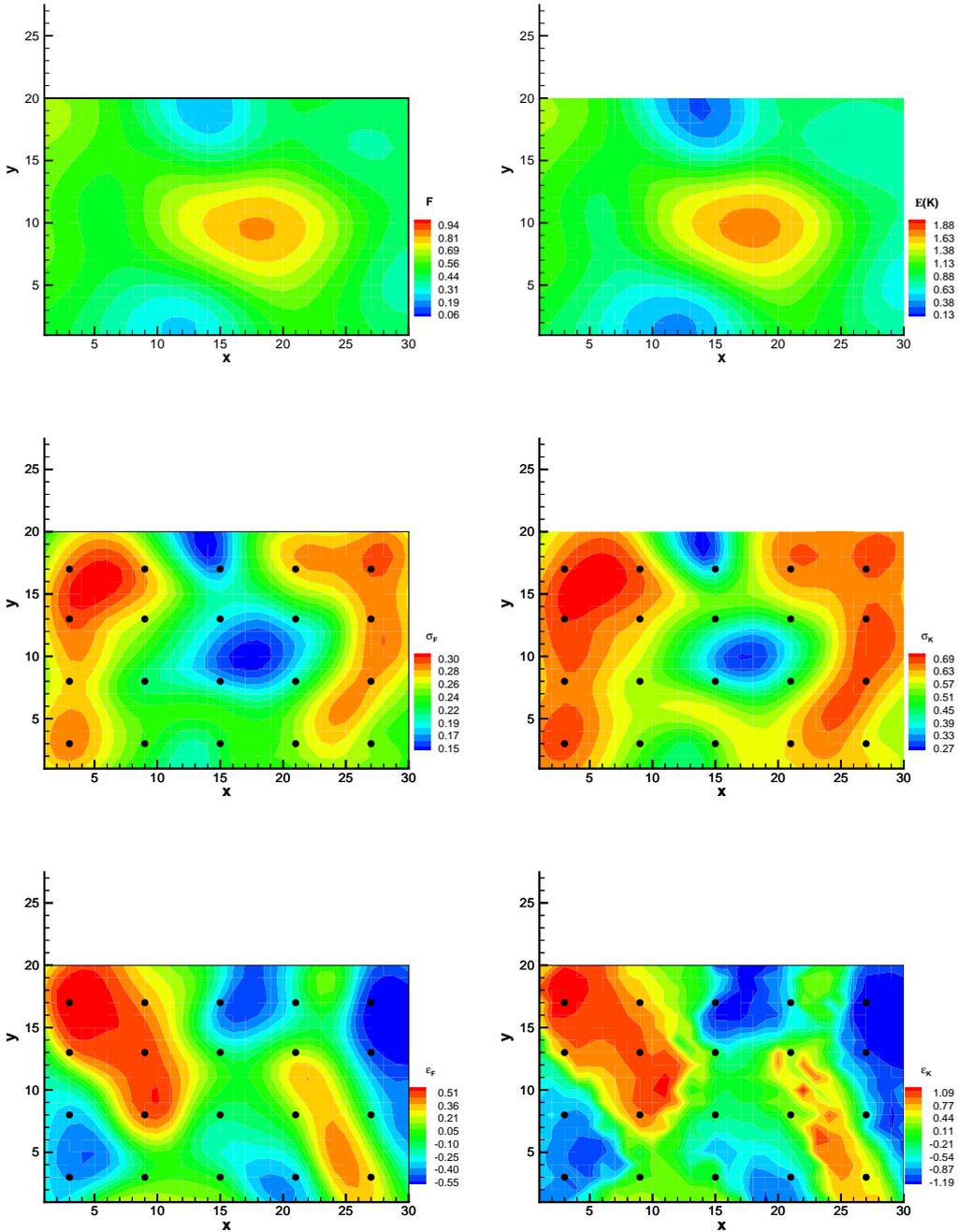


Figure 6: Results from the inference, computed using 9500 samples from the posterior distribution. Only dynamic data was used to draw the inferences. In the top row, we plot the means $\bar{\mathbf{F}}$ and $\bar{\mathbf{K}}_e$. The inferences are far worse than those seen in Fig. 4 and show little resemblance to the true $\mathbf{F}_c^t(\mathbf{x})$ and \mathbf{K}_c^t plotted as contours over them. Given such a large lack of fidelity, the standard deviations and errors plotted in the middle and last rows have little significance, but are reproduced here for completeness. Note that while Fig. 4 and Fig. 5 share the same color scales, the range had to be adjusted in this figure.

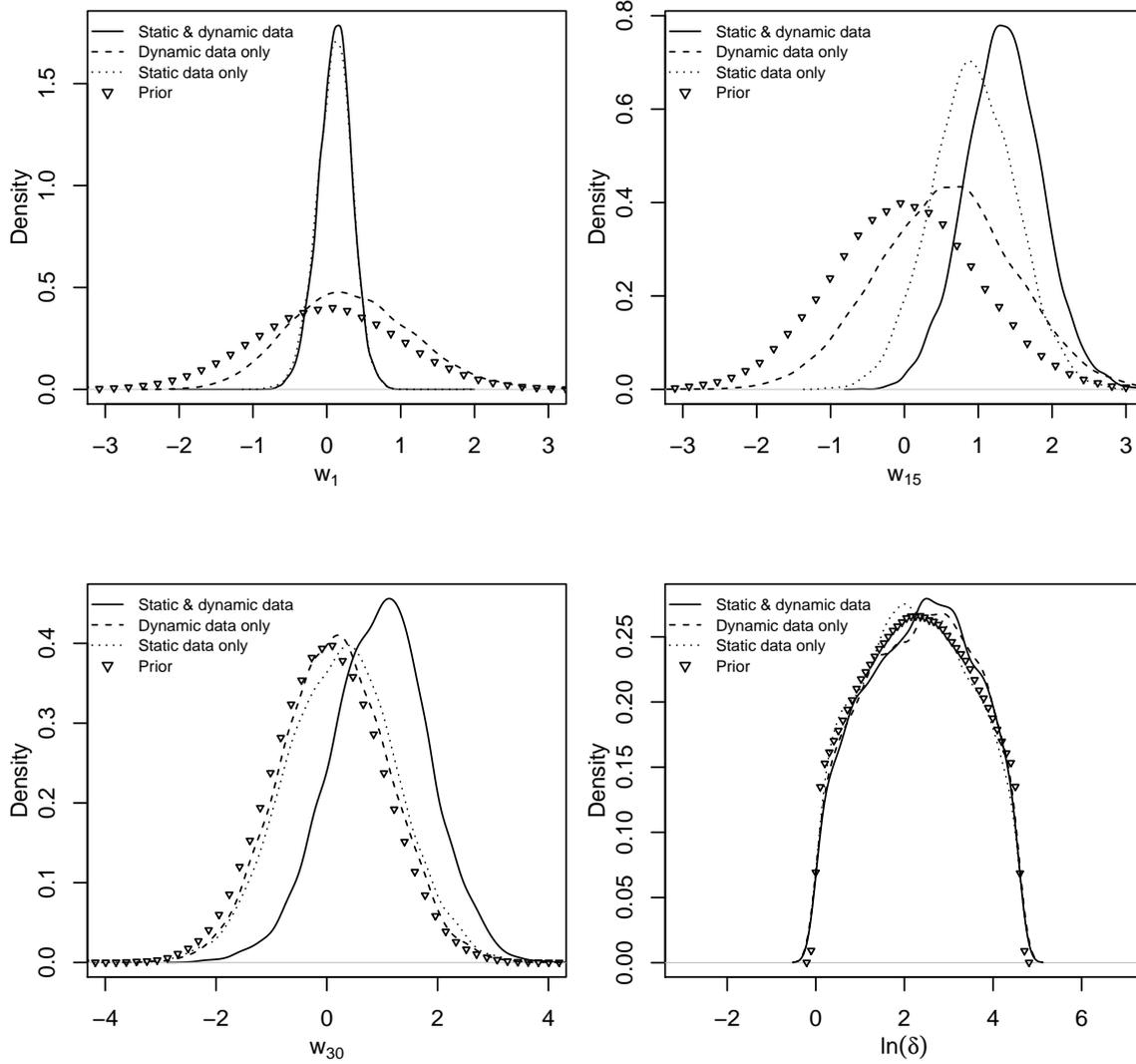


Figure 7: Marginalized PDFs of w_1, w_{15}, w_{30} and $\ln(\delta)$ as inferred from $\mathbf{k}^{(obs)}$ -only (dotted lines), $\mathbf{t}_b^{(obs)}$ -only (dashed lines) and jointly from $(\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)})$ (solid lines). The priors are plotted with ∇ for comparison. Top left: we see that the posterior for w_1 is almost entirely accounted for by the $\mathbf{k}^{(obs)}$ -only inversion (“static data only”); the posterior for w_1 , when inferred from $\mathbf{t}_b^{(obs)}$ only (“dynamic data only”) is little different from the prior. Top right: We see that both static and dynamic data contribute to the posterior of w_{15} ; the joint $(\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)})$ inversion is quite different from the other two. Bottom left: Surprisingly, neither $\mathbf{k}^{(obs)}$ nor $\mathbf{t}_b^{(obs)}$ are individually informative of the finer scales. However, when $\mathbf{k}^{(obs)}$ can account for the coarse scales, $\mathbf{t}_b^{(obs)}$ can “fill in” the information on the finer scales. This is seen in the posterior for w_{30} jointly conditioned on $(\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)})$. Bottom right: Note that the posterior density of $\ln(\delta)$ obtained from $(\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)})$ as well as just $\{\mathbf{t}_b^{(obs)}\}$ are skewed somewhat to the right of the prior, while that of the static-data-only inversion is skewed left. The prior and posteriors are truncated at $\ln(\delta) = 0, 4.6$ but the kernel density estimates used to create plots smooth them near the truncation limits.

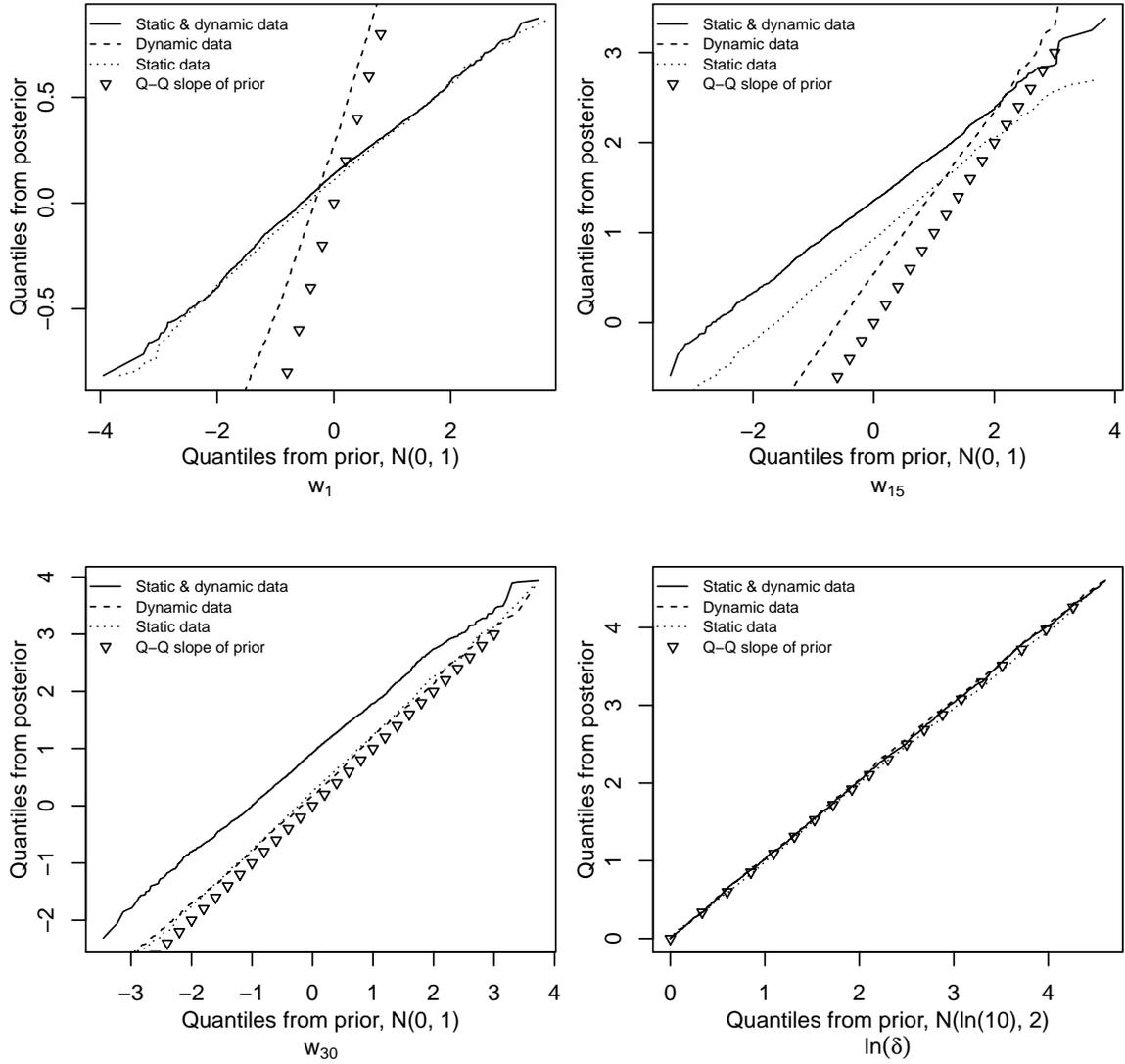


Figure 8: Top left: Q-Q plot of the posterior of w_1 developed with both static and dynamic data (solid line) compared to that obtained from static data only (dotted line) and dynamic data only (dashed line). We see that there is little difference; the length-scale of the first Karhunen-Loève mode is large enough that the sensor grid provides a sufficiently complete sampling and breakthrough times contribute little. Top right: the posteriors for w_{15} , which show the impact of dynamic data. The inference drawn from static data only underpredicts the value of w_{15} . However, the inference drawn from just the dynamic data is not very informative. Bottom left: Inference of w_{30} is determined entirely by the breakthrough times since the length-scale of the Karhunen-Loève mode is too small to be sampled by the sensors in SSA. Consequently, the posterior developed from the static data only is indistinguishable from the prior. Strangely, the posterior conditioned on just the dynamic data is indistinguishable from the prior i.e., the dynamic data, by itself, does not inform very efficiently on the finer structures. Bottom right: The posterior density for $\ln(\delta)$ obtained from $\{\mathbf{t}_b^{(obs)}\}$ is indistinguishable from that obtained with $(\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)})$.

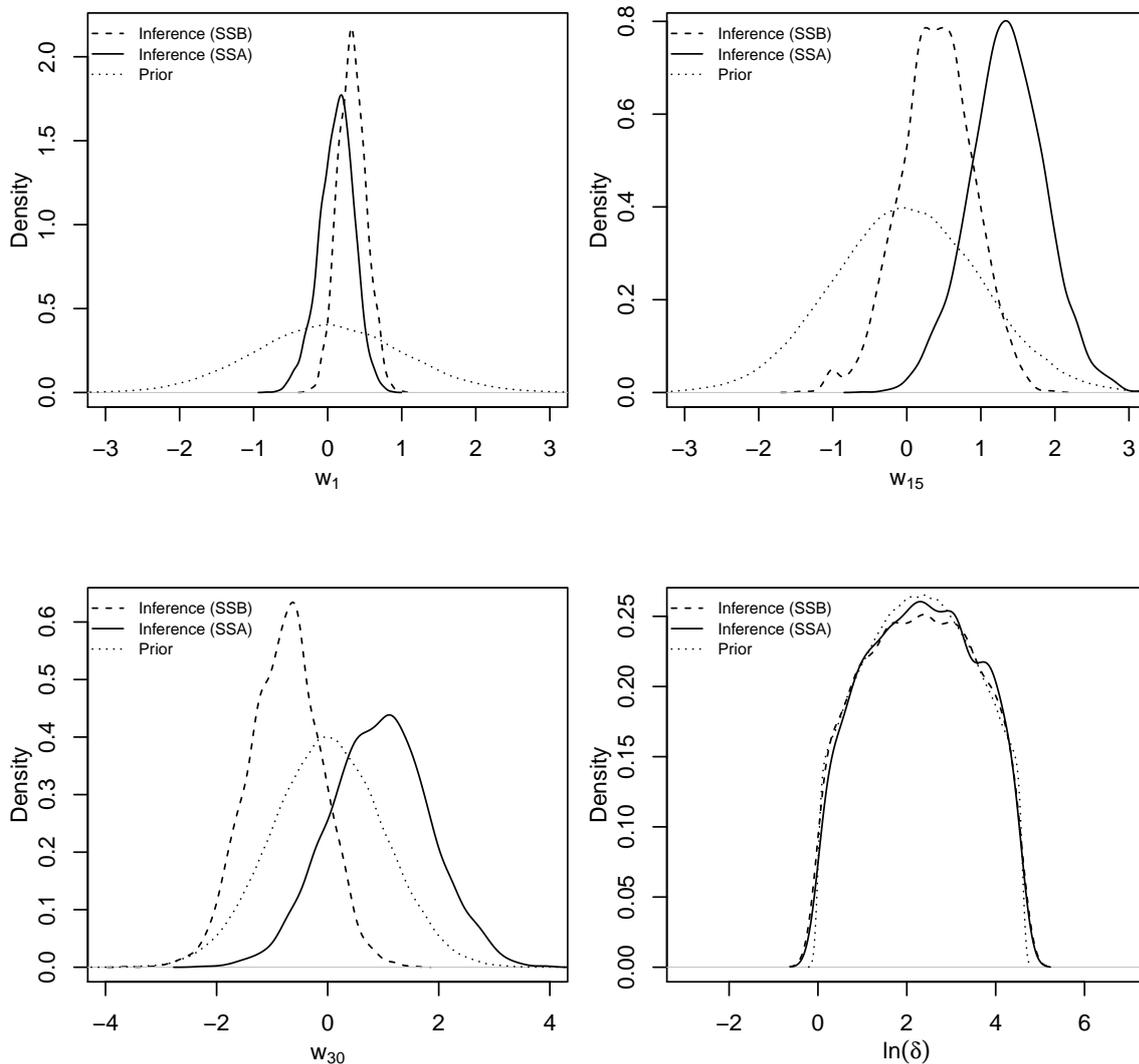
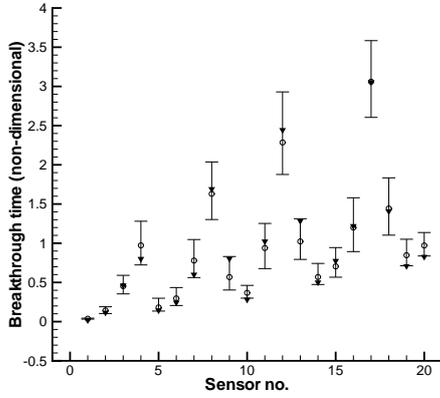
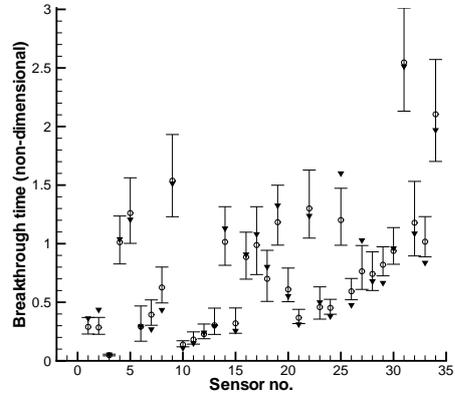
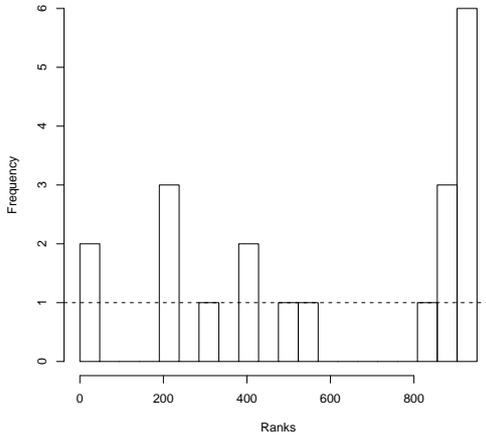


Figure 9: Prior and marginalized posterior values of the Karhunen-Loève mode weights w_1 , w_{15} , w_{30} and $\ln(\delta)$ as computed using the sensor-sets SSA (solid line) and SSB (dashed line). We see that the posteriors for w_1 are similar, indicating that the higher sensor density of SSB collects little extra information on w_1 . However, the PDFs for w_{15} and w_{30} are quite different, indicating that the uncertainty in their values may have been underestimated in both cases. The posterior density for $\ln(\delta)$ obtained from SSB sensors is shallower than that obtained from SSA, though the maximum a posteriori estimate for $\ln(\delta)$ remains approximately unchanged.



Verification rank histogram



Verification rank histogram

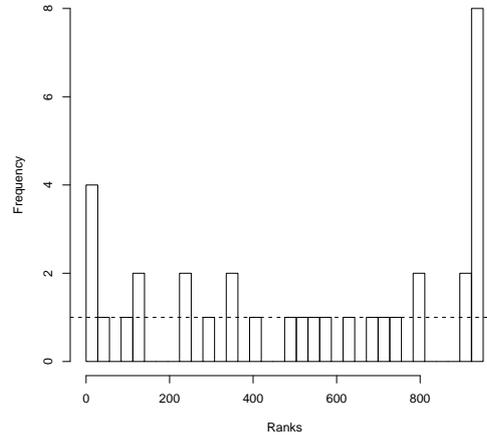


Figure 10: Results from the posterior predictive check for breakthrough times conducted by generating $\mathbf{K}_{e,i}$ and using $\mathcal{M}(\mathbf{K})$ to generate the breakthrough times. Top: The filled, inverted triangles are the observations at the sensors. The median breakthrough times from the posterior predictive checks are plotted with open symbols and the error bars denote the 1st and 99th percentiles of the breakthrough time distribution. Left: We plot the results and the observations for SSA. Right: We plot the results for SSB. We see that the difference between the 1st and 99th percentiles is generally smaller for the SSB results, indicating a reduction in predictive uncertainty. Bottom: We plot the verification rank histogram for the ranks of the observations, given the predictive distribution of breakthrough times. Ideally, the histograms should be uniform; however, the histogram for the SSB predictions is marginally better (see the deviation-from-uniformity score in Table 1).

5 Posterior predictive modeling

In this section, we explore the difference between permeability fields developed by conditioning $P(\mathbf{w}, \delta | \mathbf{d})$ jointly on static and dynamic data (multiscale data) and those reconstructed from only one type of data. We assess them based on their ability to reproduce the observed breakthrough times at the SSA sensors. We consider two flow configurations – (1) the “nominal” configuration, where the injector is in the lower left corner of the domain and (2) the “flipped” configuration, where the injector is in the upper left. The producer is always diagonally opposite to the injector. The flipped configuration is introduced to test the predictive robustness of the inferences. $P(\mathbf{w}, \delta | \mathbf{d})$ will be used to develop realizations of the *fine-scale* binary medium (on a 3000×2000 mesh) as outlined in Sec. 4.2, which will then be used in flow simulations using MODPATH.

Fig. 11 shows example realizations of the binary medium which are consistent with the multiscale observations i.e., they were developed from $\{\mathbf{w}, \delta\}$ conditioned on $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$. The white and gray regions are the high and low permeability regions, respectively. A single particle track from the injection wells to each of the SSA sensors is also shown. We choose 8 $\{\mathbf{w}, \delta\}_i$ samples from the posterior and use them to develop the corresponding $\mathbf{F}_i(\mathbf{x})$ using Eq. 18. Then, using the procedure (based on excursion sets of multiGaussian [mG] fields) described in Sec. 3.3, we develop the corresponding fine-scale binary field on a 3000×2000 mesh. The true binary field is plotted in the middle of the figure (image (e)). The variation between realizations is due to both the stochastic nature of the construction process, where white noise is convolved with a Gaussian kernel, and the sampled value of δ , which is different for each realization. Of these eight examples, realization (g) has the largest δ value and realization (i) has the smallest.

Three combinations of $\mathbf{k}^{(obs)}$ and $\mathbf{t}_b^{(obs)}$ conditioning data ($\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}, \{\mathbf{k}^{(obs)}\}, \{\mathbf{t}_b^{(obs)}\}$) were used, and for each combination, 1000 fine-scale binary realizations were created. For each of the SSA sensors and for each realization, the median travel time between the injector and the sensor is determined. Recall that sensor locations are defined at the grid-block scale and that for each location a total of 121 particles are tracked between the injector and the sensor. The median time from these 121 particles is extracted at each sensor location and designated as the breakthrough time for that location. A distribution of breakthrough times across the 1000 realizations is then created. Fig. 12 (left) shows a comparison of the three breakthrough time CDFs (cumulative distribution function) against the true breakthrough time for sensor location 14 (coordinates: 21,13). All three distributions capture the true breakthrough time (accurate) with the “Fine and Coarse” (alternatively, inferred jointly from $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$) and “Fine Only” data (alternatively, inferred solely from $\{\mathbf{t}_b^{(obs)}\}$) combinations being the narrowest (most precise) although somewhat biased towards underestimation of the true travel time. The “Coarse Only” data combination (alternatively, inferred solely from $\{\mathbf{k}^{(obs)}\}$) produces a breakthrough time distribution that is less precise but also less biased.

Fig. 12 (right) shows the precision and bias for all 20 SSA sensors and provides perspec-

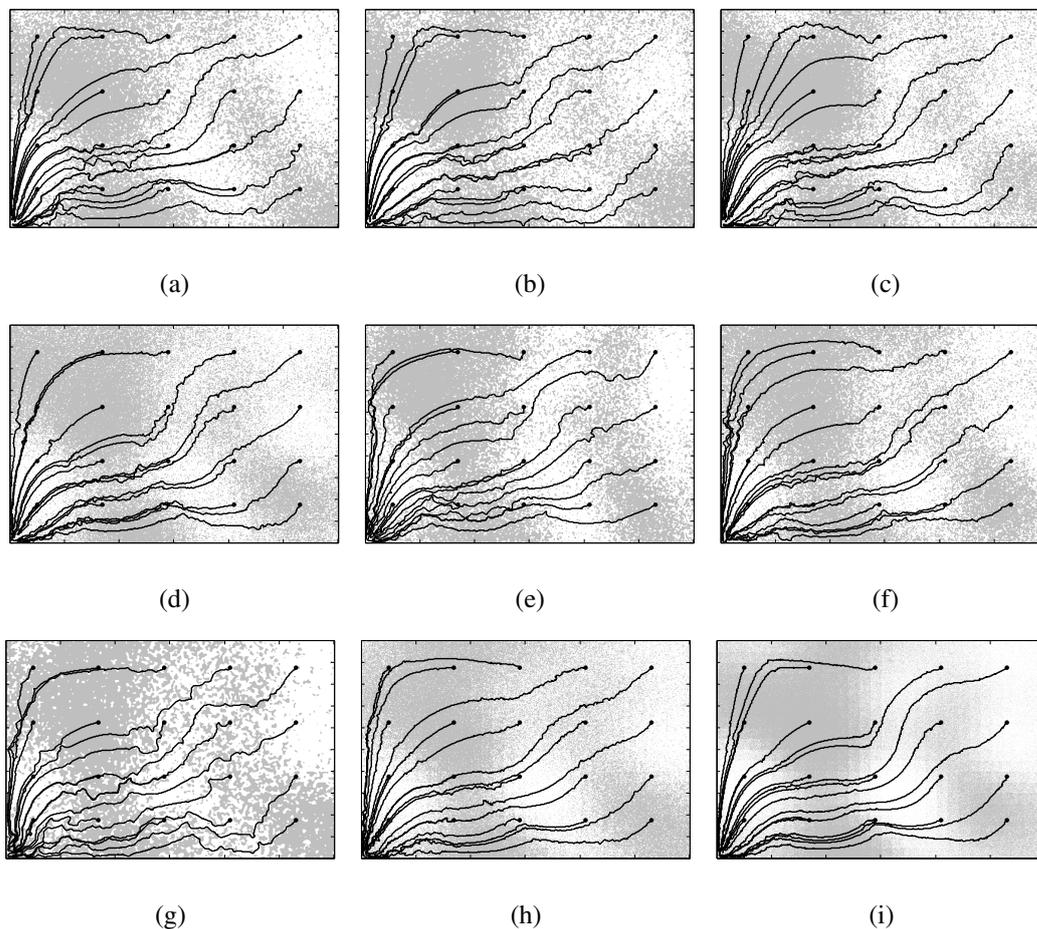


Figure 11: Fine-scale realizations of the random binary field conditioned on the observations and showing a single flowpath from the injection wells to each SSA sensor. These were developed from the posterior distribution of $\{\mathbf{w}, \delta\}$ using the mG-based technique described in Sec. 3.3. The center image (e) is the ground truth fine-scale realization. White indicates high permeability and gray is low permeability. Variation in the binary patterns are due to the stochastic nature of the process and the variation in the estimated δ parameter between realizations.

tive on how these quantities change across the model domain. For each sensor location, the black “+” indicates the location of the sensor and the relative location of the true breakthrough time. The distribution of breakthrough times from each combination of $\mathbf{k}^{(obs)}$ and $\mathbf{t}_b^{(obs)}$ is shown as a circle. The center of the circle is offset from the true breakthrough time by the amount of bias, calculated here as the difference between the median of the distribution and the true value. Underestimation shifts the center of the circle to the left and over estimation to the right. The center 95% of the breakthrough time distribution defines the radius of the circle. Both the radius and the bias offset are calculated in units of dimensionless travel time. The axes scales in Fig. 12 (right) serve as both geographic and travel time measures to define both the sensor locations and the relationship of the median breakthrough time distributions to the true breakthrough time. At each location, the order of the circles is set such that the smallest (tightest distribution) is at the front and the largest is at the back.

Examination of Fig. 12 (right) shows that all three combinations of $\mathbf{k}^{(obs)}$ and $\mathbf{t}_b^{(obs)}$ create accurate distributions of breakthrough times (all circles contain the true value denoted by the “+”). Additionally, estimated fields that are conditioned jointly on $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ create the most precise distribution (black circle is smallest and therefore on the top) for the majority of the SSA locations. Bias is relatively small at all locations as shown by all three circles being approximately centered on the true value. The spatial pattern shown Fig. 12 (right) is consistent with the flow patterns for this domain with a source and sink in the lower left and upper right corners. The breakthrough time distributions are most precise along the diagonal between the injector and producer where the majority of the flow takes place and least precise along the upper and lower boundaries where the flow fraction is quite low (see Fig. 1, b). The amount of bias is also smallest along the diagonal and greatest along the top and bottom boundaries.

In Table 2 we tabulate the bias i.e., the difference between the true breakthrough time and the median of predicted breakthrough times obtained by simulating transport through the 1000 fine-scale realizations. This is also the shift in the circles’ centers in Fig. 12. We see that the bias tends to be smallest in the $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ case where the fine-scale realizations were inferred jointly from static and dynamic observations, while the $\{\mathbf{t}_b^{(obs)}\}$ case (realizations conditioned solely on dynamic data) are the largest. This difference increases for sensors which are further way from the diagonal connecting the injector and the producer, which forms the main travel path for the fluid. Here, we introduce an additional test of the robustness of the estimated fields by “flipping” the injector-producer configuration. “Flipping” the injector-producer configuration increases the bias, most significantly for the $\{\mathbf{t}_b^{(obs)}\}$ case. In Table 3 we plot the variation in the predicted breakthrough times at each of the SSA sensors. Variation is defined as the central interval containing 95% probability mass in the CDF of breakthrough time predictions obtained from transport simulations with the 1000 fine-scale realizations. It also corresponds to the radii of the circles in Fig. 12. The table was included mainly to demonstrate the degradation of predicted breakthrough times in the “Flipped” arrangement, when realizations inferred solely from $\{\mathbf{t}_b^{(obs)}\}$ were used for predicting breakthrough times. The circle radii for certain sensors an order of magnitude

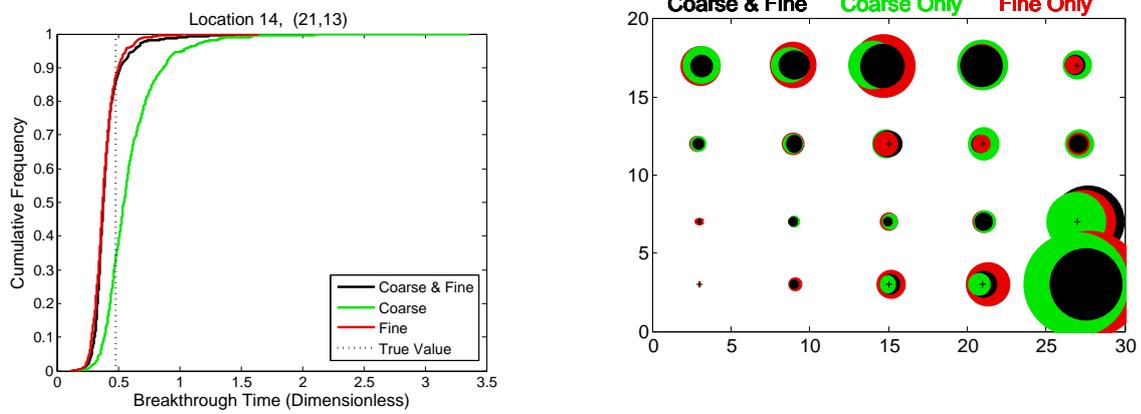


Figure 12: Comparison of the median travel times to the SSA sensors for three different data combinations. Left: Breakthrough time CDFs are shown for an example location (sensor 14). Right: The circle plots summarize the breakthrough time distributions and compare them to the true breakthrough times for all sensor locations. Details are in the text. “Coarse & fine” refer to realizations conditioned jointly on $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$, “Coarse only” refer to inferences using on $\{\mathbf{k}^{(obs)}\}$ while “Fine only” indicates inferences conditioned only on $\{\mathbf{t}_b^{(obs)}\}$.

larger than their $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ counterparts and simply could not be plotted in Fig. 12.

Fig. 12 is focused on the breakthrough time at the SSA sensor locations. However, since the ground truth in this study is known, it is possible to examine the quality of the travel time estimates for every location in the model domain. Fig. 13 (left) shows distributions of the average absolute error (AAE) between the estimated and true breakthrough time. Each average is calculated over all 600 coarse-scale grid blocks for a single realization and the distribution of the averages is then determined for all 1000 realizations. Lower values in the distribution come from realizations that better fit the true breakthrough time at all locations. Travel times are calculated for each estimated field under the nominal case injector-producer arrangement and under a “flipped” arrangement.

The majority of the fields conditioned jointly to the $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ or just the $\mathbf{k}^{(obs)}$ data have a median travel time AAE of less than 0.5 time units (Figure 13). Those fields conditioned to only $\mathbf{t}_b^{(obs)}$ data have considerably larger AAE values. The difference in results between the two different source-sink configurations is a secondary effect compared to whether or not the estimates were conditioned to $\mathbf{k}^{(obs)}$. The $\mathbf{t}_b^{(obs)}$ data alone are not enough to adequately constrain the estimations.

All comparisons shown above consider the breakthrough time values from the injector to one or more coarse grid cells. These comparisons are consistent with the conditioning data, $\mathbf{t}_b^{(obs)}$, used in the inverse estimates. However, additional information can be gained from the full distribution of travel times to each coarse grid block. The two-sample Kolmogorov-Smirnov (KS) test is used to compare the true travel time distribution to the distribution calculated on each estimated field for each coarse-scale grid block. The KS test is a non-parametric test of the difference between two distributions. The KS test statistic is the maximum vertical distance, D between two CDFs:

$$D_{i,j} = \sup_x |F_{i,n}(x) - F_{j,n'}(x)| \quad (20)$$

Here the i^{th} and j^{th} distributions have the same number of travel times, 121, for all comparisons ($n = n'$). Values of the D statistic range from 0, when there is no difference between the distributions, to 1.0 when the values of distributions do not overlap at all. The null hypothesis of the KS test is that both samples come from the same underlying population. Here, for each fine-scale realization, we compare the modeled and observed travel time distribution at every location within the model domain and calculate the test statistic, D . The average value of D across all 600 locations is retained for each of the 1000 realizations.

Fig. 13 (right) shows the resulting distributions of the average D value for six sets of 1000 realizations. The smallest values of D occur when both $\mathbf{k}^{(obs)}$ and $\mathbf{t}_b^{(obs)}$ are used as conditioning data. The distributions of D when only $\mathbf{k}^{(obs)}$ or $\mathbf{t}_b^{(obs)}$ are used have slightly larger values and are very similar. All distributions resulting from the baseline injector-producer produce lower D values than those created with the “Flipped” injector-producer configuration. The largest D values by far occur when the fine-scale realizations are conditioned solely on $\{\mathbf{t}_b^{(obs)}\}$ (fine-scale data) and then subjected to the “Flipped” conditions.

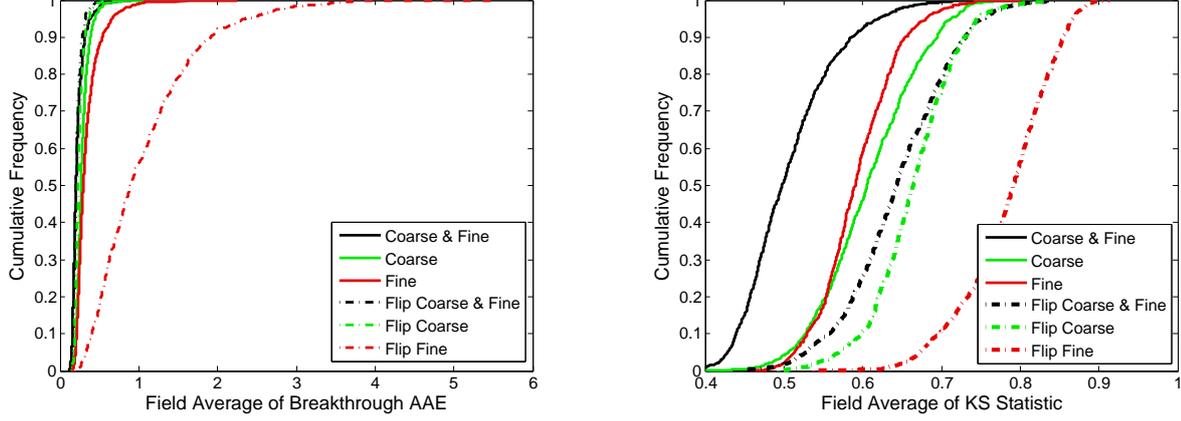


Figure 13: Average absolute errors (AAE) between median travel times from the estimated and true travel time distributions for each field (left) and distribution of average KS statistic values calculated by comparing the full distribution of estimated and true travel time distributions (right).

To summarize, we sampled the distribution $P(\mathbf{w}, \delta | \mathbf{d})$, developed in Sec. 4.3, to construct 1000 realizations of the fine-scale binary medium on a 3000×2000 mesh and predict breakthrough times at the SSA sensors via simulation. This was performed for the three different $P(\mathbf{w}, \delta | \mathbf{d})$ developed (conditioned on $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$, $\{\mathbf{k}^{(obs)}\}$ and $\{\mathbf{t}_b^{(obs)}\}$) in Sec. 4.3 to gauge the impact multiscale data in the estimation of \mathbf{K}_e . Posterior predictive model evaluations using 1000 realizations created from each of the three combinations of conditioning data clearly showed that using data collected on both scales, as opposed to a single scale, creates predictions that are closest to actual values from the ground truth. Evaluating the fine-scale fields using a flipped source-sink configuration highlights the greater impact of $\mathbf{k}^{(obs)}$ (the coarse-scale data) on the accuracy of the inference relative to using only $\mathbf{t}_b^{(obs)}$ (fine-scale data).

We now address the finding (in Sec. 4.4) that quality of the inference was mostly determined by the quality of our multiscale link model \mathcal{L} and the crude i.i.d. Gaussian error model for the discrepancy between observations of log-permeability and model predictions. In the next section, Sec. 6, we construct a more sophisticated parameterization for the model error and devise an inference scheme, predicated on the new error model and exploration of the \mathbf{K}_e space, to develop more accurate estimates of $\mathbf{F}(\mathbf{x}), \delta$ and their uncertainties.

Table 2: Tabulation of the discrepancy between the observed and predicted breakthrough times at the 20 SSA sensors. All values are in dimensionless time. The predicted breakthrough time, in this context, is the median of the breakthrough times calculated over 1000 fine-scale realizations. This result corresponds to the offset of the circles from the center in Fig. 12. The “Nominal config.” is the configuration where the fluid is injected at the lower left corner of the domain and extracted at the upper right; the “Flipped config.” is the configuration where the injector is at the top left and the producer at the bottom right. $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ indicates fine-scale realizations developed from the inversion using both static and dynamic data; “ $\{\mathbf{k}^{(obs)}\}$ ” and “ $\{\mathbf{t}_b^{(obs)}\}$ ” indicate inferences drawn from just the static (“coarse-scale”) and dynamic (“fine-scale”) observations. Note that the observations were drawn only in the nominal configuration. We see that the posterior predictions are closer to observations in the $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ case, which use both the static and dynamic data, compared to the cases when only one type of data is used. Fine-scale realizations obtained from $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ data are also surprisingly accurate when predicting breakthrough times in the “flipped” configuration. This predictive robustness is missing in the realizations drawn from only one type of data – static (coarse-scale) or dynamic (fine-scale); the errors in the predictions for the “flipped” configuration are quite large.

Sensor	Nominal config.			“Flipped” config.		
	$\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$	$\{\mathbf{k}^{(obs)}\}$	$\{\mathbf{t}_b^{(obs)}\}$	$\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$	$\{\mathbf{k}^{(obs)}\}$	$\{\mathbf{t}_b^{(obs)}\}$
(3, 3)	0.133	0.162	0.089	-0.003	-0.005	-0.006
(8, 3)	-0.044	-0.047	-0.077	-0.043	-0.051	-0.056
(13, 3)	0.000	0.003	0.002	0.088	0.093	0.196
(17, 3)	0.003	0.004	0.006	0.070	0.088	0.577
(3, 9)	0.009	-0.249	0.007	-0.038	-0.029	0.002
(8, 9)	0.037	-0.043	-0.052	-0.092	-0.080	-0.184
(13, 9)	-0.011	0.020	-0.013	0.094	0.122	0.031
(17, 9)	0.043	0.043	0.070	0.023	0.230	0.813
(3, 15)	-0.348	-0.965	-0.282	-0.185	-0.232	0.260
(8, 15)	-0.037	-0.121	-0.166	-0.117	-0.054	-0.246
(13, 15)	0.018	0.099	0.074	0.197	0.266	-0.057
(17, 15)	0.160	0.005	0.233	0.359	0.161	0.666
(3, 21)	-0.105	-0.013	-0.049	-0.171	-0.251	0.980
(8, 21)	-0.099	0.069	-0.103	-0.464	-0.322	-0.501
(13, 21)	0.058	0.105	0.068	0.312	0.356	-0.087
(17, 21)	0.061	-0.189	0.330	0.024	0.108	-0.043
(3, 27)	-0.134	-0.010	-0.202	-0.246	-0.350	2.616
(8, 27)	0.048	0.164	0.089	-0.084	-0.162	1.250
(13, 27)	0.713	-0.030	0.454	0.334	0.316	0.108
(17, 27)	0.568	-0.044	0.748	0.034	-0.094	-0.548

Table 3: Tabulation of the variation predicted breakthrough times at the 20 SSA sensors. Variation is defined as the width of the central 95% of the breakthrough times predicted by the 1,000 fine-scale realizations and corresponds to the radius of the circles in Fig. 12. The “Nominal config.” is the configuration where the fluid is injected at the lower left corner of the domain and extracted at the upper right; the “Flipped config.” is the configuration where the injector is at the top left and the producer at the bottom right. $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ indicates fine-scale realizations developed from the inversion using both static and dynamic data; “ $\{\mathbf{k}^{(obs)}\}$ ” and “ $\{\mathbf{t}_b^{(obs)}\}$ ” indicate inferences drawn from just the static (“coarse-scale”) and dynamic (“fine-scale”) observations. Note that the observations were drawn only in the nominal configuration. We see that the posterior predictions are closer to observations in the $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ case, which use both the static and dynamic data, compared to the cases when only one type of data is used. Fine-scale realizations obtained from $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ data are also surprisingly accurate when predicting breakthrough times in the “flipped” configuration. This predictive robustness is missing in the realizations drawn from only one type of data – static (coarse-scale) or dynamic (fine-scale); the errors in the predictions for the “flipped” configuration are quite large.

Sensor location	Nominal config.			“Flipped” config.		
	$\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$	$\{\mathbf{k}^{(obs)}\}$	$\{\mathbf{t}_b^{(obs)}\}$	$\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$	$\{\mathbf{k}^{(obs)}\}$	$\{\mathbf{t}_b^{(obs)}\}$
(3, 3)	0.565	1.078	1.105	0.030	0.023	0.040
(8, 3)	0.241	0.353	0.369	0.126	0.129	0.407
(13, 3)	0.047	0.057	0.122	0.267	0.270	2.444
(17, 3)	0.013	0.015	0.044	0.418	0.424	5.763
(3, 9)	0.812	1.014	1.361	0.109	0.115	0.380
(8, 9)	0.387	0.484	0.525	0.258	0.311	0.343
(13, 9)	0.159	0.226	0.200	0.718	0.762	2.138
(17, 9)	0.146	0.157	0.273	0.681	0.660	8.046
(3, 15)	1.243	1.406	1.881	0.407	0.353	1.949
(8, 15)	0.728	0.743	0.617	0.606	0.674	0.760
(13, 15)	0.188	0.363	0.438	0.734	0.897	0.830
(17, 15)	0.518	0.408	0.795	1.331	1.068	5.131
(3, 21)	1.167	1.485	1.215	0.710	0.658	5.493
(8, 21)	0.443	0.880	0.405	0.991	1.335	2.461
(13, 21)	0.440	0.563	0.594	0.789	1.108	0.550
(17, 21)	0.688	0.601	1.262	0.710	0.854	1.842
(3, 27)	0.522	0.791	0.455	1.112	1.380	14.82
(8, 27)	0.423	0.777	0.558	0.803	0.724	13.05
(13, 27)	2.173	1.741	1.881	1.018	1.505	3.674
(17, 27)	2.169	3.200	3.243	1.033	1.207	0.953

6 Multiscale inference with model errors

In Sec. 4 we demonstrated the solution of the inverse problem for inclusion distribution and size under the assumption that the discrepancy between the predictions of our models ($\mathbf{t}_b = \mathcal{M}(\mathbf{K})$ and $\mathbf{K}_e = \mathcal{L}(\mathbf{F}(\mathbf{x}), \delta)$) and the observations \mathbf{d} as Gaussians (Eq. 14). However, in most real-life cases, the discrepancy is dominated by model errors i.e., the inability of $\mathcal{M}(\mathbf{K})$ and $\mathcal{L}(\mathbf{F}(\mathbf{x}), \delta)$ to capture reality. Fig. 2 compares the predictions of effective log-permeability made by \mathcal{L} versus those computed numerically using MODFLOW-2005, for the same set of random binary media realizations. As is clear, the scatter of MODFLOW-2005 estimates around the \mathcal{L} predictions is neither symmetric nor Gaussians; imposing a normal distribution on the discrepancy, while sufficient for demonstrating an algorithm, cannot be justified when accuracy of inference is a concern. Thus, a new error model needs to be devised and used to derive a counterpart to Eq. 15. It is clear from Fig. 2 that the discrepancy in \mathbf{K}_e will be dependent on \mathbf{F} i.e., an expression for $P(\Delta\mathbf{K}|\mathbf{F})$ will have to be developed, where $\Delta\mathbf{K} = \mathbf{K}_c^t - \mathcal{L}(\mathbf{F}, \delta)$. Also, for ease of use with MCMC, a unimodal distribution for $P(\Delta\mathbf{K}|\mathbf{F})$ would be desirable.

6.1 Construction of the error model for the link function

We develop a model to evaluate $P(\Delta\mathbf{K}|\mathbf{F})$ based on kernel density estimation [71]. In Fig. 14, top left, we plot \mathbf{K}_c^t and \mathbf{K}_e as a function of \mathbf{F} , for $\delta = 38$ grid-cells. Since (\mathbf{F}, δ) is an incomplete description of the fine-scale, many fine-scale realizations, each with its own \mathbf{K}_c^t are consistent with such a specification, leading to a cloud of points (in black); the crosses indicate $\mathbf{K}_e = \mathcal{L}(\mathbf{F}, \delta)$. A joint distribution of $(\Delta\mathbf{K}, \mathbf{F})$, plotted as a scatter plot, is shown in Fig. 14, top right; this serves as our starting point for developing a kernel-smoothed model for the joint probability density function $P(\Delta\mathbf{K}, \mathbf{F})$.

We assume that the samples plotted in Fig. 14 (top right) are drawn from an unknown joint distribution which we approximate with $P(\Delta\mathbf{K}, \mathbf{F})$ as

$$\begin{aligned} P(\Delta\mathbf{K}, \mathbf{F}) &= \frac{1}{S} \sum_{k=1}^S \mathcal{K} \left(\frac{\Delta\mathbf{K} - \Delta\mathbf{K}_k}{h_{\mathbf{K}}} \right) \mathcal{K} \left(\frac{\mathbf{F} - \mathbf{F}_k}{h_{\mathbf{F}}} \right) \\ &= \frac{1}{S} \frac{1}{h_{\mathbf{K}} h_{\mathbf{F}}} \sum_{k=1}^S \mathcal{K}(\Delta\mathbf{K} - \Delta\mathbf{K}_k) \mathcal{K}(\mathbf{F} - \mathbf{F}_k) \end{aligned} \quad (21)$$

where S is the number of $(\Delta\mathbf{K}, \mathbf{F})$ samples, \mathcal{K} is an Epanetchnikov kernel and $\{h_{\mathbf{K}}, h_{\mathbf{F}}\}$ are the bandwidths in the two directions. The bandwidths are calculated using a plug-in technique [72]. The evaluation of $P(\Delta\mathbf{K}, \mathbf{F})$, per Eq. 21, involves a sum over all samples, which can be expensive; instead we employ an approximate summation method using KD-trees [73]. The samples are collated in to boxes numbering $O(\log(S))$ and stored in a KD-tree. An approximate expression for the contribution of kernels in a given box to

$P(\Delta\mathbf{K}, \mathbf{F})$ is computed and the summation is done over $O(\log(S))$ boxes. The novelty of the scheme lies in the hierarchical manner in which kernels are collated in boxes, which are further refined/sub-divided to meet an accuracy criterion. Some boxes are shown in Fig. 14, top right. The joint probability density distribution $P(\Delta\mathbf{K}, \mathbf{F})$ resulting from the kernel-smoothed approximation is shown in Fig. 14, lower left.

The evaluation of $P(\Delta\mathbf{K}|\mathbf{F})$ is performed simply by

$$P(\Delta\mathbf{K}|F) = \frac{P(\Delta\mathbf{K}, \mathbf{F})}{P(\mathbf{F})} \quad (22)$$

where $P(\mathbf{F})$ is calculated by numerically integrating out $\Delta\mathbf{K}$ in $P(\Delta\mathbf{K}, \mathbf{F})$. Computations involving KDE were performed using the MATLABTM code in [74]. Fig. 14, lower right shows $P(\Delta\mathbf{K}|\mathbf{F})$ for different values of \mathbf{F} .

Note that the bandwidths $\{h_{\mathbf{K}}, h_{\mathbf{F}}\}$ obtained using plug-in bandwidth estimation had to be over-smoothed by a factor of 2 to ensure that $P(\Delta\mathbf{K}|F)$ did not have largely disconnected modes (since they prevent MCMC chains from mixing).

In order to construct a model for $P(\Delta\mathbf{K}|\mathbf{F})$, valid for arbitrary \mathbf{F} and δ , the exercise was repeated for $\delta = 9, 15, 21, 28, 38, 56$ and 75 grid-cells. Models for $P(\Delta\mathbf{K}|\mathbf{F}, \delta)$ were developed for each. Thereafter, $P(\Delta\mathbf{K}|\mathbf{F}, \delta)$ for arbitrary values of δ were obtained by adopting that of the nearest known δ .

Kernel density estimates of probability density distributions are generally applicable only in the region where samples exist and can be misleading away from it. Our use of Epanetchnikov kernels (which have compact support) ensures that at large $\Delta\mathbf{K}$, $P(\Delta\mathbf{K}|\mathbf{F}, \delta)$ will evaluate to zero, in keeping with the absence of samples. This will be reflected when devising an expression for $P(\mathbf{d}|\mathbf{K}_e, \mathbf{F}, \delta)$ in Eq. 15 i.e., the expression will evaluate to zero in large parts of the multi-dimensional (\mathbf{w}, δ) space, as will the posterior distribution. This can pose a problem when exploring the space of the posterior density using a MCMC sampler - large regions of zero gradient in the posterior will not allow the chain to move to more promising regions. We ameliorate this situation by adding an extra Gaussian kernel \mathcal{K}_G to Eq. 21, i.e.

$$P(\Delta\mathbf{K}, \mathbf{F}) = \frac{1 - \alpha}{S} \sum_{k=1}^S \mathcal{K} \left(\frac{\Delta\mathbf{K} - \Delta\mathbf{K}_k}{h_{\mathbf{K}}} \right) \mathcal{K} \left(\frac{\mathbf{F} - \mathbf{F}_k}{h_{\mathbf{F}}} \right) + \alpha \mathcal{K}_G(\Delta\mathbf{K} - \mu_k, \Gamma_k) \quad (23)$$

where α is chosen so that only 10% of the probability mass is contributed by the Gaussian kernel. The infinite support of the Gaussian allows the posterior to be calculated anywhere in the (\mathbf{w}, δ) space and provides the gradient required by the MCMC chain to move into the high probability regions of the posterior. In the regions in (\mathbf{w}, δ) space where $\Delta\mathbf{K}$ is small, the kernel density estimate for the model errors' probability density dominate, allowing the inference to reflect their effect (uncertainty due to model errors), rather than the measurement noise.

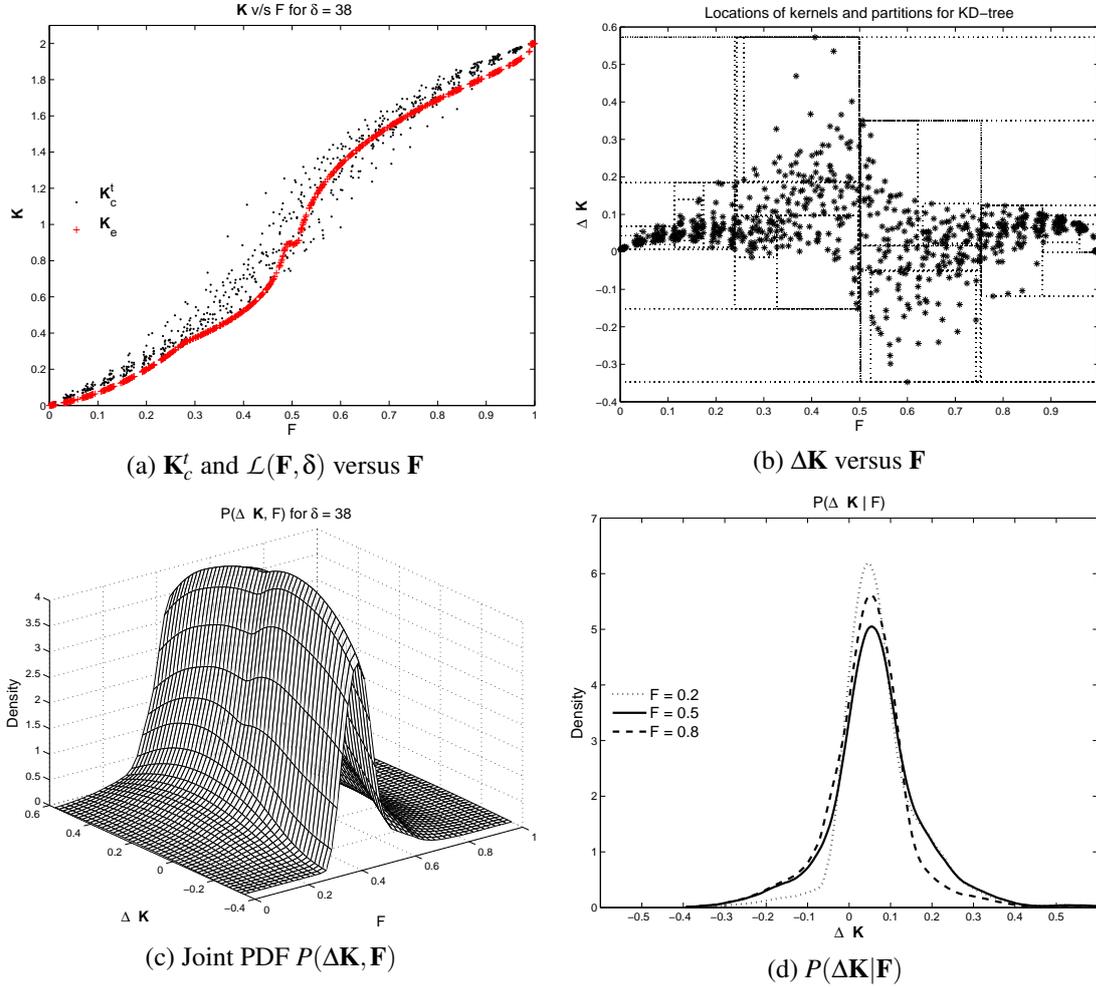


Figure 14: Stages in the construction of the conditional PDF $P(\Delta \mathbf{K} | \mathbf{F}, \delta)$. Top left: We create multiple realizations of the fine-scale random binary media field and plot their true upscaled log-permeability (as a cloud) and the model prediction $\mathbf{K}_e = \mathcal{L}(\mathbf{F}, \delta)$, for $\delta = 38$ grid-cells. Top right: we show the samples and the boxes from Level 4 of the KD-tree holding the boxed samples. Bottom left: We plot the kernel density estimate of $P(\Delta \mathbf{K}, \mathbf{F})$ for $\delta = 38$ grid-cells, constructed using Epanetchnikov kernels (after over-smoothing the bandwidths to remove isolated modes). Bottom right: We plot the conditional distribution $P(\Delta \mathbf{K} | \mathbf{F})$, for $\delta = 38$ grid-cells, for various values of \mathbf{F} .

6.2 Adapting and solving the inverse problem for model errors

In this section, we incorporate the expression for $P(\Delta\mathbf{K}|\mathbf{F})$ (Eq. 22), but calculated from the “ameliorated” joint density distribution in Eq. 23, into a formulation of an inverse problem (similar to Eq. 15) and solve it to infer fine-scale properties.

We start with the assumption that the measurement errors are far smaller than the model error / discrepancy $\Delta\mathbf{K}$ as defined above. In such a case

$$\mathbf{k}^{(obs)} \approx \mathbf{K}_c^t = \mathcal{L}(\mathbf{F}(\mathbf{x}), \delta) + \Delta\mathbf{K}.$$

Let $k_j^{(obs)}$ be the measured log-permeability in the grid-box containing sensor j and $\mathbf{F}(\mathbf{x}_j)$ be the corresponding inclusion proportion. Assuming that the model errors, conditioned on the local $\mathbf{F}(\mathbf{x}_j)$ are i.i.d,

$$P(\mathbf{k}^{(obs)}|\mathbf{F}(\mathbf{x}), \delta) = \prod_{j=1}^{N_s} P(k_j^{(obs)}|\mathbf{F}(\mathbf{x}_j), \delta) = \prod_{j=1}^{N_s} P(\Delta\mathbf{K}_j|\mathbf{w}, \delta) \quad (24)$$

where $\Delta\mathbf{K}_j$ is the observation-prediction mismatch of log-permeabilities at sensor j (but now attributed to model errors rather than measurement errors, as was done in Eq. 15). We model the mismatch between observed and predicted breakthrough times as before, i.e., as measurement errors modeled using i.i.d. Gaussians, implying that the model $\mathcal{M}(\mathbf{K})$ is an accurate representation of tracer transport. Thus the counterpart to Eq. 16, incorporating model errors, is

$$P(\mathbf{w}, \delta|\mathbf{d}) \propto \exp\left(-\frac{[\mathbf{e}_t(\mathbf{w}) - \boldsymbol{\mu}_t]^T [\mathbf{e}_t(\mathbf{w}) - \boldsymbol{\mu}_t]}{\sigma_t^2}\right) \prod_{j=1}^{N_s} P(\Delta\mathbf{K}_j|\mathbf{w}, \delta) \pi(\delta) \prod_{l=1}^M \exp(-w_l^2) \quad (25)$$

Eq. 25 is solved using the same techniques and for the same problem as Sec. 4. However, we restrict ourselves to the SSA set of sensors.

The new model for observation-prediction discrepancy in $\Delta\mathbf{K}$ will allow the MCMC chain to visit regions of the log-permeability (\mathbf{K}) space that may not have been allowed by the Gaussian model used in Sec. 4. However, this exploration will also be modified by the Gaussian error model used for \mathbf{t}_b . We examine the effect of these error models on $\{\mathbf{w}, \delta\}$ individually.

In Fig. 15, we plot the posterior PDFs for w_1 and w_{30} , as inferred solely from static data, as Q-Q plots. Posterior distributions, as obtained in Sec. 4, are also plotted. We see that the KDE error model constructed in Sec. 6.1 makes a difference only for the low mode (w_1) whereas the high mode (w_{30}) is unaffected. This is in keeping with the conclusions drawn

in Sec. 4. Note, however, that the posterior PDF for w_1 , obtained with the KDE model is different from the one obtained with the Gaussian model of Sec. 4; for instance, the Q-Q plot is not a straight line and so the posterior is no longer a Gaussian.

We now investigate the effect of \mathbf{t}_b by incorporating dynamic data into our inference. In Fig. 16 we plot the posterior PDFs for the Karhunen-Loève weights w_1, w_{15}, w_{30} and $\ln(\delta)$ as computed using model errors. The plots with solid lines were inferred using both static and dynamic data, whereas the ones with dotted lines were inferred using static data only. We see that like our results in Sec. 4 the static data determines the large-scale structures - the posteriors for w_1 obtained from the static observations only is very similar to the one obtained from static and dynamic observations i.e. the inclusion of \mathbf{t}_b contributes very little information. On the other hand, as described in the previous paragraph the posterior and prior densities for w_{30} in the static-observations-only case are indistinguishable (Fig. 15) and the posterior in the lower left sub-figure of Fig. 16 is due almost entirely to \mathbf{t}_b observations. $\ln(\delta)$ seems very little affected by the inclusion of dynamic data as well as the model errors. We also plot the posterior PDFs for the same variables as obtained in Sec. 4 (i.e., from Fig. 9, but only for the SSA set of sensors). We see a substantial difference in the posterior PDF for the larger modes, but the higher/finer modes are very similar. Thus the effect of KDE model for errors is most felt by inference variables that have global effect e.g. w_1 , (and are impacted by static data) whereas those with local effect are dominated by the information content of the breakthrough times (where the error model is the same as in Sec. 4). By the same argument, if the inference were to be done with static data only, the posterior density for w_{30} would be the same as the prior (and completely unaffected by the change in the error model between Sec. 4 and this section), whereas the largest change would be seen in w_1 . This is corroborated in Fig. 15. Also, comparing the width of the PDFs for the objects of inference, we do not see much of decrease in width, i.e., while the posterior distributions are different for all the objects of inference, there has not been a reduction of uncertainty. However, whether or not there is an improvement in the goodness of fit (i.e., in the accuracy of the inference) will be evaluated next.

We perform posterior predictive checks (PPC) using the posterior distribution of $\{\mathbf{w}, \delta\}$. The distribution of breakthrough times obtained at the 20 sensors of the SSA set are plotted in Fig. 17 (the “error bars” correspond to the 1st and 99th percentiles of the 9,500 breakthrough times obtained using $\{\mathbf{w}, \delta\}$ samples from the posterior distribution). The observed breakthrough times are plotted as triangles whereas the circles are the median of the PPC. The corresponding values from Sec. 4 are plotted in red. We see that the KDE error model did not contribute to the reduction of uncertainty (as mentioned in the last paragraph) since the widths of the error bars obtained with the KDE error model and Sec. 4’s Gaussian model are about the same; however, the error bars are slightly shifted and only 2 sensor readings (as opposed to 3 in Sec. 4.2) are outside the bounds. This indicates a slight improvement in the goodness-of-fit.

To summarize, we developed a KDE-based representation for the model error and performed the inference of $\{\mathbf{w}, \delta\}$ based on the assumption that model errors dominated measurement errors. We obtained inferences which were different from those obtained in

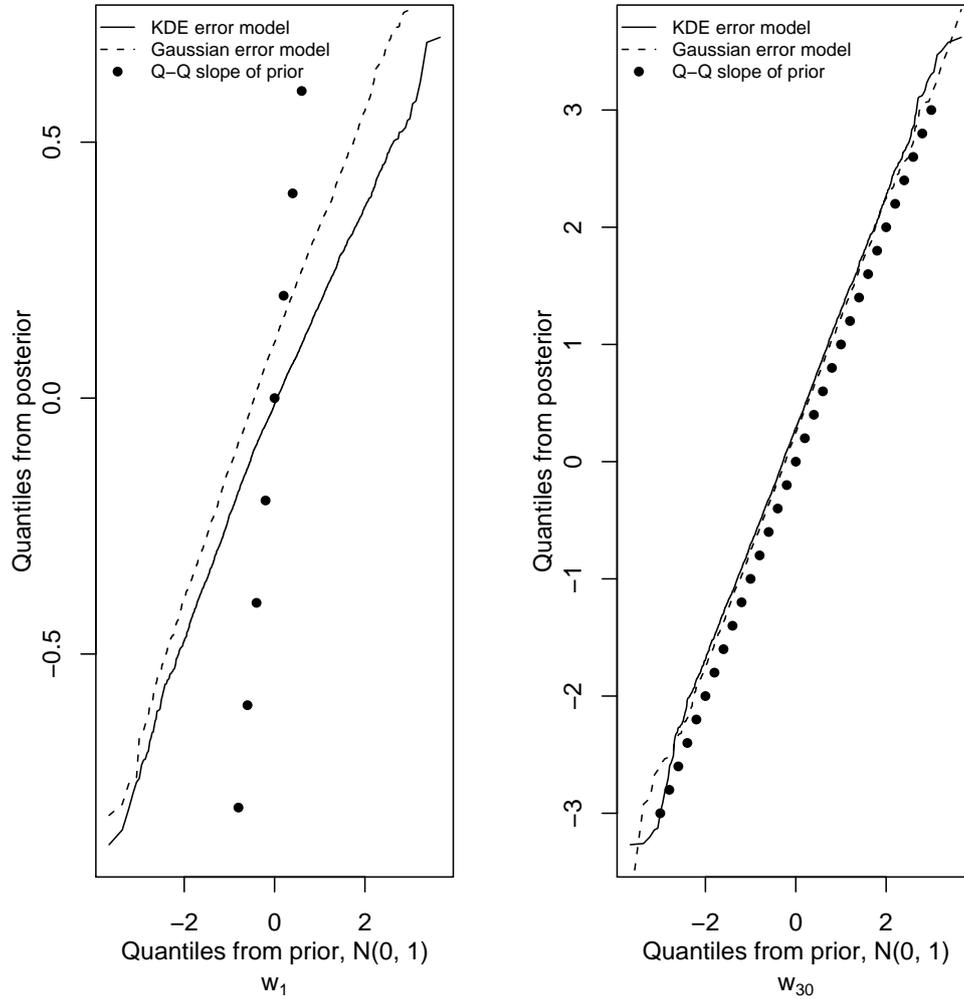


Figure 15: Q-Q plot for the posterior distribution of w_1 (left) and w_{30} (right). The solid line denote the Q-Q plot for inference developed with the KDE error model described in this section; the dashed line denote those that used the Gaussian model of Sec. 4. Only static data was used. The prior is plotted with dots. We note that the KDE model results in a plot that is not a straight line (for w_1); also the plots are different when the KDE and Gaussian models are used. Right, we see that the static data made no difference to the inference of w_{30} ; the posterior is indistinguishable from the prior.

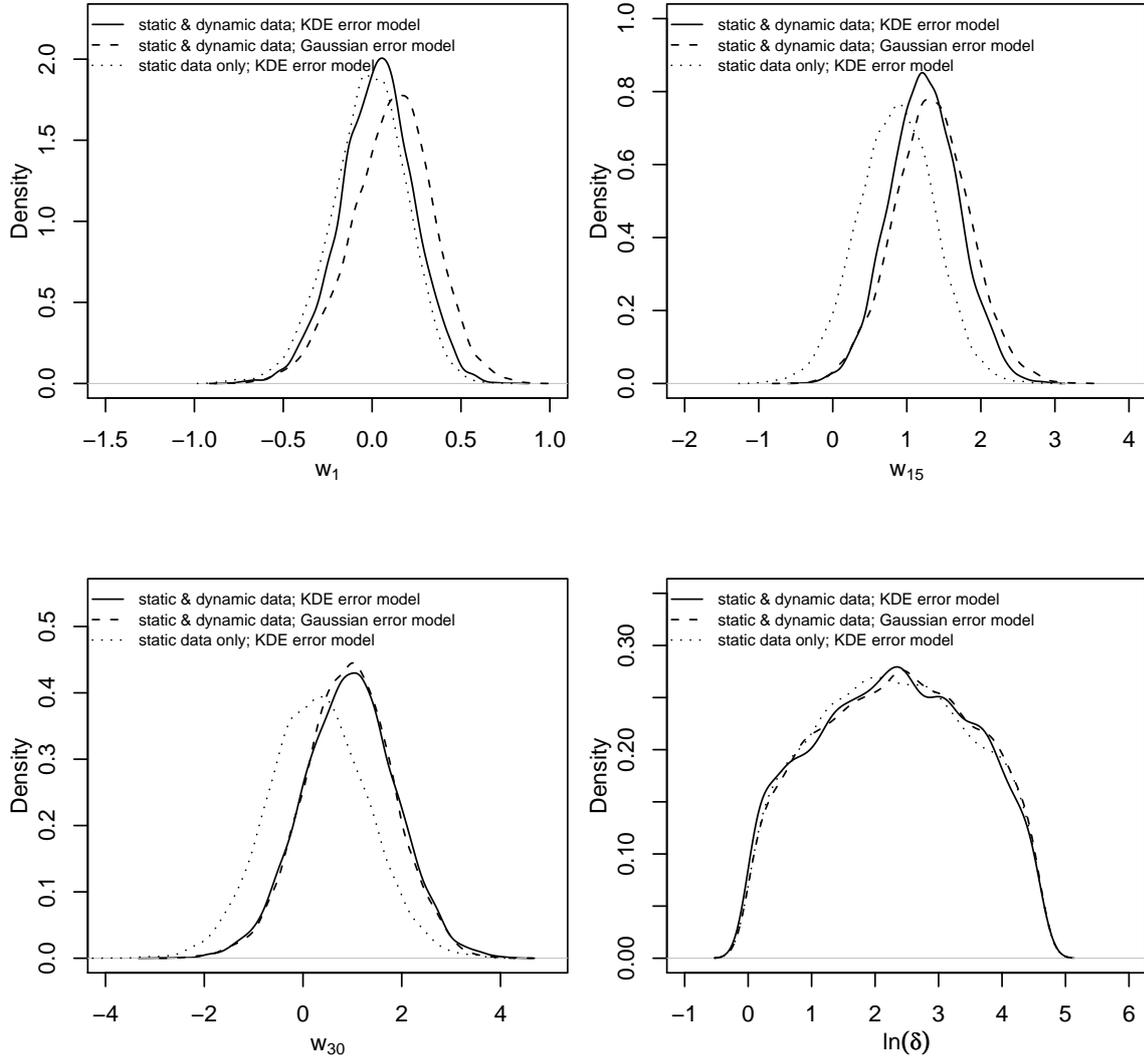


Figure 16: Posterior PDFs for $\{w_1, w_{15}, w_{30}, \ln(\delta)\}$ using the KDE error model, obtained with static data only (dotted lines), as well as static and dynamic data (solid line). We see that dynamic data affects the posterior for the higher Karhunen-Loève modes, as in Sec. 4. Also the impact of the static data is limited to the lower Karhunen-Loève modes. We see that the impact of dynamic data on $\ln(\delta)$ is low. We also plot the posterior PDF obtained using the Gaussian error model in Sec. 4. We see that at the finer (higher) Karhunen-Loève modes, which are impacted mostly by dynamic data, have posteriors that are relatively insensitive to the choice of KDE versus Gaussian data model.

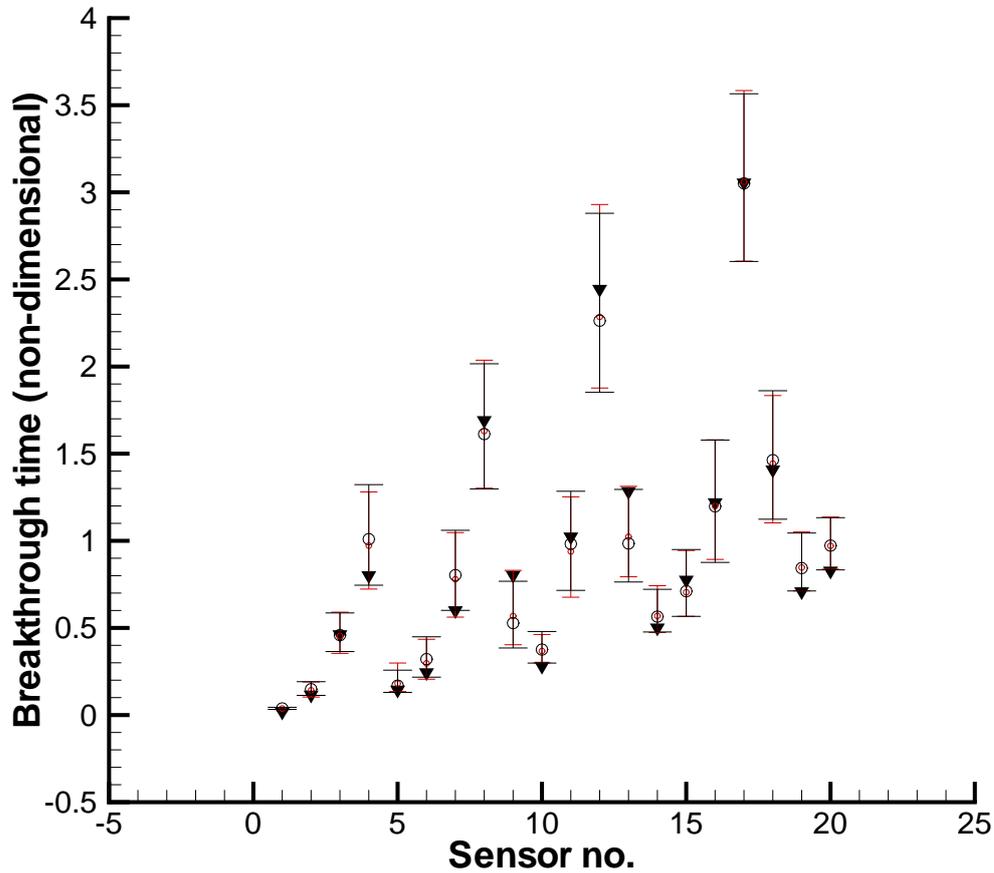


Figure 17: Distributions of the tracer breakthrough times at the 20 SSA sensors, obtained via posterior predictive checks using the posterior density for $\{\mathbf{w}, \delta\}$. Those plotted in black were developed using the KDE error model; the ones in red use the Gaussian model described in Sec. 4. The observations are in black triangles; the circles are medians. The error bars indicate the 1st and 99th percentiles. We see that the width of the error bars are about the same irrespective of the error model (KDE/Gaussian) used; i.e. our choice of models did not reduce the uncertainty in the inference. However, only 2 observations fall outside the black error bars, whereas 3 do for the red ones, indicating a slight improvement in the goodness-of-fit when using the KDE error model.

Sec. 4, but these were restricted to variables which had global effect; the posterior for higher Karhunen-Loève modes, which are controlled by the breakthrough times, were hardly affected. A local variable like $\ln(\delta)$ also was unaffected. The reason for the change was the ability of the MCMC chain to explore parts of the \mathbf{K} space which was not possible under the Gaussian error model in Sec. 4. While we did obtain a better fit of the model to data, as indicated by the PPC, it was marginal. The reason for this lies in our restrictive model for $\mathbf{K}_e = \mathcal{L}(\mathbf{F}(\mathbf{x}), \delta)$; the KDE model for $\Delta\mathbf{K}$ shows that the \mathbf{K} -space is quite involved and may not be easily captured by the Karhunen-Loève -mode based reduced model. This inability to capture the spatial complexity of \mathbf{K} also results in poor predictions for \mathbf{t}_b .

A solution to this problem of the rigidity/limited range of \mathcal{L} lies in actually using the KDE model for $\Delta\mathbf{K}$, along with \mathcal{L} , to construct proposals for \mathbf{K}_c^t . These \mathbf{K}_c^t could be compared with observations $\mathbf{k}^{(obs)}$ (under a measurement error assumption) and also used to generate \mathbf{t}_b and compared to $\mathbf{t}_b^{(obs)}$, again under a tight measurement error model. This would undoubtedly improve the fit of the model to data, but will also result in a far more difficult inference problem - apart from inferring $\{\mathbf{w}, \delta\}$, we would also have to infer a \mathbf{K}_c^t field under a suitable smoothness model. This, much harder, inference problem will be addressed in a subsequent publication.

This page intentionally left blank

7 Conclusions

We have developed a multiscale, statistical technique to reconstruct characteristics of a random, porous binary medium from partial observations. The binary medium consists of high permeability inclusions distributed in a low-permeability matrix. The uneven spatial distribution of the inclusions $\mathbf{F}(\mathbf{x})$ and its size δ are the objects of inference; effective permeability $\mathbf{K}_e = \mathcal{L}(\mathbf{F}(\mathbf{x}), \delta)$ is calculated from a known dependence on them. The inclusions are too small to be resolved on a mesh. The observations consist of measurements of log-permeability ($\mathbf{k}^{(obs)}$) and breakthrough times ($\mathbf{t}_b^{(obs)}$) of a fluid from a pump test at a set of sensor points. The log-permeability measurements inform on the large-scale variations in the domain of interest; the breakthrough times are governed strongly by the contorted flow paths through the medium and are informative on the fine-scale structures. Jointly conditioning on these observations renders this a *multiscale* inverse problem.

The reconstruction is posed as a Bayesian inverse problem, predicated on fitting a multiscale transport model to the data. The transport model is formulated at the coarse-scale, but with a statistical subgrid model that incorporates the impact of the fine, unresolved scales at the coarse-scale. The subgrid model (alternatively, the *link* function between the scales) is parameterized with the inclusion size, δ . The inverse problem is regularized by expressing $\mathbf{F}(\mathbf{x})$ using Gaussian Processes; its dimensionality is reduced by expanding $\mathbf{F}(\mathbf{x})$ in terms of a truncated Karhunen-Loève series. Fitting the model to data yields a joint distribution of the Karhunen-Loève weights and δ , inferring characteristics of both the coarse-scale (the variation of $\mathbf{F}(\mathbf{x})$ and \mathbf{K}_e in the domain) and the fine, unresolved scale (the size δ of the inclusions). This distribution is realized without any approximations, using an adaptive MCMC sampler.

We find that $\mathbf{F}(\mathbf{x})$ and δ obtained by jointly conditioning on $\{\mathbf{k}^{(obs)}, \mathbf{t}_b^{(obs)}\}$ are far more accurate than if they were obtained solely from $\mathbf{k}^{(obs)}$ or $\mathbf{t}_b^{(obs)}$. Inversion based on $\mathbf{k}^{(obs)}$ capture the large-scale variation correctly; however, the permeability field distribution so reconstructed predicts breakthrough times poorly since it lacks finer details/structures responsible for contorted flow paths. The inversions based solely on $\mathbf{t}_b^{(obs)}$ are extremely poor, due to its inability to constrain large structures. However, when the two are put together, $\mathbf{k}^{(obs)}$ constrains the larger structures while $\mathbf{t}_b^{(obs)}$ is used to constrain the smaller ones. Thus the improved inferences (when conditioned jointly on $\mathbf{k}^{(obs)}$ and $\mathbf{t}_b^{(obs)}$) are not just due to more plentiful observations, but rather due to the different *types* of information in $\mathbf{k}^{(obs)}$ and $\mathbf{t}_b^{(obs)}$. Since the information is derived from different scales (coarse for $\mathbf{k}^{(obs)}$ and fine for $\mathbf{t}_b^{(obs)}$), the inversion is fundamentally multiscale.

Plots of marginalized posteriors of the objects of inference show that they are approximately Gaussian. While the priors used for them are Gaussian, this is surprising, given that both the fluid transport and the link function are strongly nonlinear. However, this (approximate) Gaussian character of the posterior raises the possibility of performing the inference accurately using scalable methods like ensemble Kalman filters which maintain accuracy

only up to the second moment of the distribution that they infer.

We use the inferences of $\mathbf{F}(\mathbf{x})$ and δ to reconstruct fine-scale realizations of the binary medium i.e., on a fine mesh where the inclusions can be resolved. We check the predictive skill of the ensemble of realizations by performing transport simulations with them, in an effort to recreate the observations from which they were inferred in the first place. We find that fine-scale realization developed by conditioning jointly on coarse and fine-scale observations (i.e., “multiscale” realizations) have a greater predictive skill than those inferred with one type of information. Further, the predictive skill of the multiscale realizations are preserved when they are used in flow configurations different from the one used to obtain $\mathbf{t}_b^{(obs)}$. This robustness is not observed in the realizations obtained from only one type of data.

We find that the primary source of uncertainty and inaccuracy in multiscale inversions is the approximation inherent in the link function, which in turn leads to a significant model error when predicting permeability using $\mathcal{L}(\mathbf{F}(\mathbf{x}), \delta)$. We characterize this model error (an upscaling error) statistically, using a mixture of Epanechnikov kernels (a kernel density estimate between true and \mathcal{L} predictions of permeability) and incorporate into the inversion. Our results indicate that this sophisticated, heteroscedastic error model (as opposed to using a homoscedastic model for upscaling error) improves the quality of the inference. This was verified using a posterior predictive test.

This work opens up a few new questions. While we find that multiscale inversion leads to more robust inferences of binary media, the Bayesian technique is computationally expensive and will probably scale only to problems of modest size. Approximating the posterior distribution as a Gaussian opens up the possibility of employing scalable Ensemble Kalman Filtering techniques, which is currently being investigated. A second avenue of research lies in whether the fine-scale realizations constructed from the data can be used in an efficient manner for predictions; currently, ensemble predictions require many simulations on a fine mesh. While an embarrassingly parallel problem, it nevertheless consumes significant computing resources and often does not provide a timely solution. Finally, we consider whether the incorporation of model errors into the inversion process can be performed in a more accurate manner. We are currently exploring ways in which the model errors can be used to jointly infer the Karhunen-Loève weights, δ and a permeability field from the same observations.

This page intentionally left blank

References

- [1] D. S. Oliver and Y. Chen. Recent progress on reservoir history matching: a review. *Computational Geosciences*, 15:185–221, 2011.
- [2] H. J. Hendricks Franssen, A. Alcolea, M. Riva, M. Bakr, N. van der Wiel, F. Stauffer, and A. Guadagnini. A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments. *Advances in Water Resources*, 32:851–872, 2009.
- [3] H. Lee, D. Higdon, Z. Bi, M. Ferreira, and M. West. Markov random field models of high-dimensional parameters in simulations of fluid flow in porous media. *Technometrics*, 44(3):230–241, 2002.
- [4] J. Wang and N. Zabarvas. A Markov random field model of contamination source identification in porous media flow. *International Journal of Heat and Mass Transfer*, 49:939–950, 2006.
- [5] Jianlin Fu and J. Jaime Gómez-Hernández. A blocking Markov chain Monte Carlo method for inverse stochastic hydrogeological modeling. *Mathematical Geosciences*, 41:105–128, 2009.
- [6] Jianlin Fu and J. Jaime Gómez-Hernández. Uncertainty assessment and data worth in groundwater flow and mass transport modeling using a blocking Markov chain Monte Carlo method. *Journal of Hydrology*, 364:328–341, 2009.
- [7] W. Li and O. A. Cirpka. Efficient geostatistical inverse methods for structured and unstructured grids. *Water Resources Research*, 42, 2006. W06402.
- [8] P. K. Kitanidis. Quasi-linear geostatistical theory for inversing. *Water Resources Research*, 31(10):2411–2419, 1995.
- [9] Youssef M. Marzouk and Habib N. Najm. Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems. *Journal of Computational Physics*, 228(6):1862 – 1902, 2009.
- [10] B. Jafarpour and D. McLaughlin. Efficient permeability parameterization with discrete cosine transform. In *The Proceedings of the SPE Reservoir Simulation Symposium*, number Paper SPE 106453, 2007.
- [11] B. Jafarpour and D. B. McLaughlin. History matching with an ensemble Kalman filter and discrete cosine parameterization. *Computational Geosciences*, 12:227–244, 2008.
- [12] B. Jafarpour, V. K. Goyal, D. B. McLaughlin, and W. T. Freeman. Compressed history matching: Exploiting transform-domain sparsity for regularization of nonlinear dynamic data integration problems. *Mathematical Geosciences*, 42:1–27, 2010.

- [13] M. A. R. Ferreira, Z. Bi, M. West, H. K. H. Lee, and D. Higdon. Multi-scale modeling of 1-D permeability fields. volume 7, pages 519–527. Oxford University Press, 2003.
- [14] S. I. Aanonsen and D. Eydinov. A multiscale method for distributed parameter estimation with application to reservoir history matching. *Computational Geosciences*, 10:97–117, 2006.
- [15] D. Higdon, H. Lee, and Z. Bi. A Bayesian approach to characterizing uncertainty in inverse problems using coarse and fine scale information. *IEEE Transactions in Signal Processing*, 50(2):389–399, 2002.
- [16] Y. Efendiev, T. Hou, and W. Luo. Preconditioning Markov chain Monte Carlo simulations using coarse-scale models. *SIAM Journal for Scientific Computing*, 28(2):776–803, 2006.
- [17] Y. Efendiev and T. Hou. Multiscale finite element methods for porous media flows and their applications. *Applied Numerical Mathematics*, 57(5-7):577–596, 2007. Special Issue for the International Conference on Scientific Computing.
- [18] M. Hayek, F. Lehmann, and P. Ackerer. Adaptive multi-scale parameterization for one-dimensional flow in unsaturated porous media. *Advances in Water Resources*, 31:28–43, 2008.
- [19] I. Berre, M. Lien, and T. Mannseth. A level-set corrector to an adaptive multiscale permeability prediction. *Computational Geosciences*, 11:27–42, 2007.
- [20] L. K. Nielsen, H. Li, X-C Tai, S. I. Aanonsen, and M. Espedal. Reservoir description using a binary level set method. *Computing and Visualization in Science*, 13:41–58, 2010.
- [21] I. Berre, M. Lien, and T. Mannseth. Multilevel parameter structure identification for two-phase porous-media flow problems using flexible representations. *Advances in Water Resources*, 32:1777–1788, 2009.
- [22] M. Sahimi. *Flow and transport in porous media and fractured rock: From classical methods to modern approaches*. Wiley-VCH, 1995.
- [23] K. S. Mendelsohn. A theorem on the effective conductivity of a two-dimensional heterogeneous medium. *Journal of Applied Physics*, 46(11):4740–4741, 1975.
- [24] A. H. Sihvola and J. A. Kong. Effective permittivity of dielectric mixtures. *IEEE Transactions on Geosciences and Remote Sensing*, 26(4):420–429, 1988.
- [25] G. Dagan. Models of groundwater flow in statistically homogeneous porous formations. *Water Resources Research*, 15(1):47–63, 1979.
- [26] G. Dagan. *Flow and transport in porous media*. Springer-Verlag, 1989.
- [27] A. D. Poley. Effective permeability and dispersion in locally heterogeneous aquifers. *Water Resources Research*, 24(11):1921–1926, 1988.

- [28] S. Pozdniakov and C. F. Tsang. A self-consistent approach for calculating the effective hydraulic conductivity of a binary, heterogeneous medium. *Water Resources Research*, 40, 2004. doi:10.1029/2003WR002617. W05105.
- [29] T. Miloh and Y. Benveniste. A generalized self-consistent method for the effective conductivity of composites with ellipsoidal inclusions and cracked bodies. *Journal of Applied Physics*, 63(3):789–796, 1988.
- [30] T. Miloh and Y. Benveniste. On the effective conductivity of composites with ellipsoidal inhomogeneities and highly conducting interfaces. *Proceedings of the Royal Society of London, A*, 455:2687–2706, 1999.
- [31] R. W. Zimmerman. Effective conductivity of a two-dimensional medium containing elliptical inhomogeneities. *Proceedings of the Royal Society*, (452):A1713–A1727, 1996.
- [32] P. Renard and G. de Marsily. Calculating equivalent permeability: A review. *Advances in Water Resources*, 20(5-6):253–78, 1997.
- [33] C. Knudby, J. Carrera, J. D. Bumgardner, and G. E. Fogg. Binary upscaling: The role of connectivity and a new formula. *Advances in Water Resources*, 29:590–604, 2006.
- [34] A. J. Desbarats. Numerical estimation of effective permeability in sand-shale formations. *Water Resources Research*, 23(2):273–286, 1987.
- [35] A. R. Solow. Mapping by simple indicator kriging. *Mathematical Geology*, 18(3):335–352, 1986.
- [36] A. G. Journel and F. Alaber. Non-Gaussian data expansion in the earth sciences. *Terra Nova*, 1(2):123–134, 1989.
- [37] J. J. Gómez-Hernández and R. M. Srivastava. ISIM3D: An ANSI-C three-dimensional multiple indicator conditional simulation program. *Computers and Geosciences*, 16(4):395–440, 1990.
- [38] *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, 1998.
- [39] S. F. Carle and G. E. Fogg. Transition probability-based indicator geostatistics. *Mathematical Geology*, 28(4):453–476, 1996.
- [40] S. F. Carle and G. E. Fogg. Modeling spatial variability with one and multidimensional continuous lag markov chains. *Mathematical Geology*, 29(7):891–918, 1997.
- [41] C. Lantuejoul. *Geostatistical Simulation: Models and Algorithms*. Springer-Verlag, 2002.
- [42] M. Armstrong, A. G. Galli, H. Beucher, G. Loch, D. Renard, B. Doligez, R. Eschard, and F. Geffroy. *Plurigaussian Simulations in Geosciences*. Springer, 2003.

- [43] R. J. Adler. On excursion sets, tube formulas and maxima of random fields. *Annals of Applied Probability*, 10(1):1–74, 2000.
- [44] R. J. Adler, J. E. Taylor, and K. J. Worsley. *Applications of Random Fields and Geometry: Foundations and Case Studies*. 2009.
- [45] J. E. Taylor and R. J. Adler. Euler characteristics for Gaussian fields on manifolds. *The Annals of Probability*, 31(2):533–563, 2003.
- [46] K. J. Friston, K. J. Worsley, R. S. J. Frackowiak, J. C. Mazziotta, and A. C. Evan. Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*, 1:210–220, 1994.
- [47] K. J. Worsley, S. Marrett, P. Neelan, and A. C. Evans. Searching scale space for activation in PET images. *Human Brain Mapping*, 4(1):74–90, 1996.
- [48] D. J. Nott and R. J. Wilson. Multi-phase image modelling with excursion sets. *Signal Processing*, 80:125–139, 2000.
- [49] F. M. Phillips and J. L. Wilson. An approach to estimating hydraulic conductivity spatial correlation scales using geological characteristics. *Water Resources Research*, 25(1):141–143, 1989.
- [50] S. A. McKenna, J. Ray, Y. Marzouk, and B. van Bloemen Waanders. Truncated multiGaussian fields and effective conductance of binary media. *Advances in Water Resources*, 34:617–626, 2011.
- [51] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors. *Markov Chain Monte Carlo in Practice*. Chapman & Hall, 1996.
- [52] H. Haario, E. Saksman, and J. Tamminen. An adaptive Metropolis algorithm. *Bernoulli*, 7:223–242, 2001.
- [53] Heikki Haario, Marko Laine, Antoinietta Mira, and Eero Saksman. DRAM-Efficient adaptive MCMC. *Statistics and Computing*, 16(4):339–354, 2006.
- [54] J. Goodman and A. D. Sokal. Multigrid Monte-Carlo method: conceptual foundations. *Physical Review D*, 40:2035–2071, 1989.
- [55] H. Haario, E. Saksman, and J. Tamminen. Adaptive proposal distribution for random-walk Metropolis algorithm. *Computational Statistics*, 14:375–395, 1999.
- [56] L. Tierney and A. Mira. Some adaptive Monte Carlo techniques for Bayesian inference. *Statistics in Medicine*, 18:2507–2515, 1999.
- [57] P. J. Green and A. Mira. Delayed rejection in reversible jump Metropolis-Hastings. *Biometrika*, 88:1035–1053, 2001.
- [58] A. Mira. Ordering and improving the performance of Monte Carlo Markov Chains. *Statistical Science*, 16:340–350, 2002.

- [59] R. J. Adler and J. E. Taylor. *Random fields and geometry*. Springer, 2007.
- [60] M. Griogoriu. *Stochastic calculus*. Birkhauser, 2002.
- [61] R. Ghanem and P. Spanos. *Stochastic finite elements: A spectral approach*. Springer-Verlag, 1991.
- [62] Jorg E. Aarnes, Tore Gimse, and Knut-Andreas Lie. *Geometric modelling, numerical simulation, and optimization*, chapter An introduction to the numerics of flow in porous media using Matlab, pages 265–306. Springer, 2007.
- [63] J.B. Bell and P. Colella. A second-order projection method for the incompressible Navier-Stokes equations. *J. Comp. Phys.*, 85:257–283, 1989.
- [64] A. W. Harbaugh. MODFLOW-2005, the U.S. Geological Survey modular ground-water model – the ground water flow process. U.S. Geological Survey Techniques and Methods 6-A16, U.S. Geological Survey, 2005.
- [65] D. W. Pollack. User’s guide for modpath/modpath-plot, version 3: A particle tracking post-processing package for modflow, the u.s. geological survey finite-difference ground-water flow model. U.S. Geological Survey Open-File Report 94-464, U.S. Geological Survey, 1994.
- [66] A. E. Raftery and S. M. Lewis. The number of iterations, convergence diagnostics and generic metropolis algorithms. In W. R. Gilks, D. J. Spiegelhalter, and S. Richardson, editors, *Practical Markov Chain Monte Carlo*. Chapman and Hall, 1995.
- [67] Gregory R. Warnes. *mcgibbsit: Warnes and Raftery’s MCGibbsit MCMC diagnostic*, 2005. R package version 1.0.5.
- [68] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. ISBN 3-900051-07-0.
- [69] Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007.
- [70] Tilmann Gneiting, Fadoua Balabdaoui, and Adrian E. Raftery. Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(2):243–268, 2007.
- [71] David W. Scott. *Multivariate density estimation*. Wiley Series in Probability and Statistics, 1992.
- [72] P. Hall, S. J. Sheather, M. C. Jones, and J. S. Marron. On optimal data-based bandwidth selection in kernel density estimation. *Biometrika*, 78(2):263–269, 1991.
- [73] A. G. Gray and A. W. Moore. Very fast multivariate kernel density estimation via computational geometry. In *Proceedings of the Joint Statistical Meeting*, 2003.

[74] Alexander Ihler. Kernel Density Estimation Toolbox in MATLAB (R13). <http://www.ics.uci.edu/~ihler/code/kde.html>, 2003.

DISTRIBUTION:

1	Jaideep Ray, 08954	MS 9159
1	S. A. McKenna, 06911	MS 0751
1	B. van Bloemen Waanders, 01442	MS 1318
1	Technical Library, 08944 (electronic)	MS 0899